Chapter 3 The Teleological Theory of Representation

3.1 Introduction

In this chapter and the next I shall be considering two topics which are widely regarded as raising difficulties for physicalism. This chapter will be concerned with mental representation. The next chapter will deal with consciousness. It is not difficult to see why mental representation is often thought to present a problem for physicalism. Mental states like beliefs, desires, hopes, fears, and the other propositional attitudes have representational contents: they represent the world as being a certain way. But how can this be, if such mental states involve nothing more than physical states of the brain? If my belief that Lima is the capital of Peru is realized by an arrangement of neurones, then how does this belief manage to reach out across the world and latch on to a city I have never seen? How can a bank of neurones be about something outside my head?[1]

Different ph ysicalist theories of mind, such as functionalism, or Davidsonian anomalous monism, or any of the many other physicalist accounts of mind currently on offer, will make this problem precise in rather different ways. However, since my aim in this chap ter is to defend a positive solution -- the teleological theory of representation -- which will be available to physicalists of all kinds, it will not matter greatly exactly which version of physicalism we start with. So I shall follow the pattern o f much recent literature, and start once more with functionalism.

The overall plan of this chapter will be as follows. In the next section (3.2) I shall show how repesentation arises as a problem for functionalism, and offer the tel eological theory of representation as an initial solution. Then, after some brief comments about broad propositional attitudes (3.3), I shall elaborate some of the details of the teleological theory, in the course of answering the standard objection that some beliefs serve biological purposes even when they are false (3.4). This will prompt some discussion of the status of belief-desire psychology (3.5), and also show how the teleological theory incorporates, rather than competes with, the ide a that truth guarantees the satisfaction of desires (3.6). Sections 3.7-10 will then defend this satisfaction-guaranteeing component in the teleological theory against a number of objections, and will also consider some alternative theories which sh are this satisfaction-guaranteeing assumption, but do not incorporate it within a teleological context. After this I shall return to the issue of broad beliefs, showing how it is unsurprising, given the teleological theory, that beliefs and desiress hould fail to supervene on brain states (3.11-12). The final two sections of the chapter will then discuss the availability of empirical evidence for the teleological theory (3.13), and point out the radically anti-verificationist implications of th e theory (3.14).

3.2 Functionalism and Representation

Functionalism views beliefs and desires and other mental states as internal causal intermediaries between perception and behaviour. For functionalism, w e might say, beliefs and desires are part of a system of internal pushes and pulls which explains why people behave as they do. This functionalist picture of mental states raises immediate questions about representation. After all, why should components in an internal causal structure be credited with representational powers? Surely an internal

causal role is one thing, and a representational relationship to an (almost invariably) extra-cranial state of affairs another. Functionali sm seems to describe only the first, causal aspect of mental states, and to omit the second, representational aspect. As it is sometimes put, functionalism seems to give us only the "syntax" of mental states, and to leave out their "semantics".

It is true that most versions of functionalism follow everyday practice and identify beliefs and desires in terms of "content clauses", as the belief that p, the desire that q, and so on. However, from the perspective of the rest of the functionalist package, this need only be viewed as the most convenient among many possible ways of indicating the causal structure of beliefs and desires, as one way of "labelling" causal roles, and not as an essential use of representational notions.&nb sp; After all, how could representational relationshipships to often distant states of affairs be intrinsic to the internal causal roles of mental states?

It is perhaps worth pausing on this point. Despite what I have just said, doe sn't the functionalist approach to the mind need to invoke assumptions about what desires are for and beliefs are about, in order to infer what agents will do? Well, functionalism does indeed attend to the causal roles of mental states; and, a s I have just said, it does take these causal roles to be indexed by content clauses. But, to repeat, it is not essential to this that the content clauses specify what beliefs are about or desires are for. A nice way to bring this out is to th ink of contents, as some philosophers do, in terms of sets of possible worlds. On this account, the content of an instrumental belief that F will cause G is the set of worlds in which F does cause G, and the content of a desire for G is the set of w orlds in which G obtains. Given this, and given that agents tend to perform those actions that they believe are necessary for what they want, functionalism could then invoke, as a first aproximation, the generalization that an agent will do F just i n case the set of worlds which comprises the content of the agent's desires is contained in the set of worlds which comprises the content of the agent's instrumental beliefs about F. Note, however, that it does not matter to this generalization that these beliefs and desires represent the world as being a certain way -- that they are true (in the case of beliefs) or satisfied (desires) just in case the actual world is a member of the set of worlds which constitutes the content. All the general ization needs are the overall sets of worlds which comprise the contents, since these alone suffice to specify the interdependent causal roles of beliefs and desires; it is irrelevant that these contents also determine, together with the actual worl d, whether beliefs are true or desires satisfied. Which is why, from the functionalist point of view, any other similar structures could in principle serve to specify causal roles instead, even if they didn't involve the entities we normally think o f beliefs and desires as about -- provided, that is, that they at least succeed in tying mental states to the bits of behaviour, the Fs, which are the end points of the causal roles the functionalist is interested in.

So the complaint is that functionalism gives only internal causal roles, and not representation. It might seem to some readers, however, that the difficulty is easily remedied. Isn't the trouble just that functionalism thinks of the "inputs" and "outputs" of caus al roles too narrowly, with inputs starting with the sense organs, and outputs finishing with bodily movements? So why not simply extend our causal net to allow more distal causes of perception, on the input side, and more distal effects of behaviou r, on the output side? This would allow us to analyse the truth conditions of beliefs as those distal circumstances which cause them, and the satisfaction conditions of

desires as those distal states of affairs they give rise to, and would the reby seem to reintroduce aboutness without further ado.

This move, however, is fatally afflicted by the disease known as "disjunctivitis".[2]   The belief that there is an ice-cream in front of you can be caused, not on ly by a real ice-cream, but also by a plastic ice-cream, or a hologram of an ice-cream, or so on. So, on the current suggestion, the belief in question ought to represent either-a-real-ice-cream-or-a-plastic-one-or-any-of-the-other-things-that-might -fool-you. Which of course it doesn't.

Similarly with desires. The results which follow any given desire include not only the real object of the desire, but also various unintended consequences. So the current suggestion would imply that the object of any desire is the disjunction of its real object with all those unintended consequences. Which of course it isn't.

So, even if we widen functionalism's causal roles to include distal causes and effect s, we still need somehow to winnow out, from the various causes that give rise to beliefs, and the various results that eventuate from desires, those which the beliefs are about, and which the desires are for.

This is where an appeal to teleological considerations seems to yield a natural and satisfying answer. We can pick out a desire's real satisfaction condition as that effect which it is the desire's biological purpose to produce. And, similarly, we can pick out the real truth condition of a belief as that condition which it is the biological purpose of the belief to be co-present with.[3]

This teleological theory of representation will be elaborated and defended in detail in what follo ws. But at this stage let me make two immediate points. First, my use of "purpose" and similar phrases should be understood, as in chapter 2, in terms of the aetiological account of teleological notions. That is, I take it that the purpo se of A is to do B just in case A is now present because in the past some selection process selected items that do B. So, in the specific context at hand, when I speak of that condition which it a desire's biological purpose to produce, I take it th at some past selection mechanism has favoured that desire -- or, more precisely, the ability to form that type of desire -- in virtue of that desire producing that effect. And when I speak of the condition which it is the bio logical purpose of a belief to be co-present with, I take it that some past selection mechanism has selected that belief -- or, more precisely, the ability to form that belief type -- in virtue of its occurring in conjunction with that condition. (As in chapter 2, those readers who dislike the aetiological analysis of purposive talk can simply replace all my references to purposes by references to selection mechanisms. What matters to my story is that mental states shou ld be the products of selection processes, not what terminology we use to specify this.)

The second immediate point I wish to make is that this selectionist-teleological approach to mental representation does not imply that all representa tional abilities must be genetically innate products of inter-generational selection. For selection-based teleology can also be a product of individual learning (cf. "The pigeon is pressing the bar in order to get food"). And so, if some non-i nnate belief or desire is selected in the course of individual learning in virtue of the condition it is co-present with, or the result it gives rise to, then that belief or desire will have a genuine selection-based representational purpose, despite its non-innateness.

3.3 Broad Contents

It will be helpful, before proceding to further details of the teleological theory, to comment briefly on the relation between my argument so far and the recent debate about "br oad" versus "narrow contents". My reference, in the middle of the last section, to the possibility of "widening" the functionalist net, may have made some readers think of recent philosophical discussions of "broad contents". However, our curr ent concerns are rather more general than the debate about broad contents. Our present topic is to understand content as such: why do mental states, of whatever kind, have contents? The debate about broad contents, by contrast, takes the existence of contents as such for granted, and is concerned with more detailed questions about which specific mental states have which specific contents.

The debate about broad contents arises from the observation, to which I drew attent ion in chapter 1, that the content of many beliefs seems to depend, not just on the believer's physical make-up, but also on features of the context. Thus Hilary Putnam has argued that the identity of beliefs about natural kinds depends on what kind s are actually present in the believer's world (1975); Tyler Burge has argued that the contents of theoretical beliefs can depend on features of the social context (1979, 1982); and Gareth Evans has maintained that the possession of singular b eliefs demands the existence of the objects those beliefs are about (1982).

These philosophers and their followers form one side of the debate about broad contents. On the other side is a sizeable minority who are suspicious of broa d beliefs, on the grounds, roughly, that it is hard to see how differences which lie outside the head can matter to the explanatory significance of mental states (cf Fodor, 1987).

This is why I said the debate about broad contents is less general than our current concerns. The participants in the debate take it as given that our beliefs and other attitudes have representational contents. The point of dispute is only whether or not these contents are fixed by internal physical make-up.

John McDowell (1986, sect 5) has suggested that it is only possible to get worried about the general possibility of representation if you make the mistake of thinking that all beliefs are narrow. McDowell's thought is that a problem about representation only arises as long as we think of beliefs as things inside people's head. Once we recognize that the very possession of a belief can involve extra-cranial facts, we ought no longer to be puzzled about how things insid e the head can stand for things outside.

This seems to me to get things exactly the wrong way round. Merely accepting that the possession of beliefs involves entities outside believers' heads does little to explain how representatio n as such is possible. After all, plenty of the other states that people possess involve entities outside their heads -- for example, financial solvency, or popularity, or being married -- without thereby becoming representat ional.

Far from appealing to broad contents to dissolve the general problem of representation, I think we will do better to solve the general problem of representation first, and then apply the solution to the issue of broad contents.&nbs p; In the absence of any general understanding of representation as such, much of the current debate between the friends and enemies of broad contents has collapsed into an indecisive

trading of intuitions. However, once we have arrived at a satisfa ctory general theory of content, then we shall understand why it is quite unsurprising that some contents should be broad. I shall return to this issue in section 3.10 below.

3.4 Functional Falsity

A good w ay to develop the details of the teleological theory is to consider a familiar objection. This is the objection that certain beliefs have biological purposes which require them to be present when they are false, and so constitute prima facie counter - examples to the teleological thesis that truth conditions can be analysed as those circumstances in which beliefs are biologically supposed to be present. (Ned Block has urged this objection on me. See also Stich, 1982, p 53.)

 ; For example, consider the belief that you are not going to be injured in some unavoidable and imminent trial of violence. It is arguable that natural selection has bequeathed us an innate disposition to form this belief, even in cases where it hig hly likely that we will in fact be injured, in order to ensure that we will not flinch in battle. But it then seems to follow that, according to the teleological theory, the truth condition of this belief will include many cases where we will be inj ured -- since such cases will be among those where we are biologically supposed to have the belief. So we seem to have a reductio of the teleological theory. For by hypothesis the truth condition of the belief is that we won't be i njured.

 Examples like this are interesting, but I don't think they suffice to discredit the teleological theory. In order to see why not, we need to consider the way that beliefs and desires combine to generate actions in the overal l human decision-making system. It will emerge that the purpose of beliefs in this system is to guide actions in such a way that desires will be satisfied. And then, by understanding the teleological theory as focusing on this specific purpose of beliefs, we will be able to accommodate examples of the above kind. The point will be that stopping you flinching is a special kind of biological purpose, which cuts across the purpose of satisfying desires,[4] and which therefore does not require the truth of beliefs in the way that the satisfaction of desires does.

 The overall biological function of the human decision-making system is to generate actions that cause biologically suitable results. Beliefs an d desires both contribute to this purpose. However, they contribute in different ways. The role of desires is to do with the fact that different results are suitable at different times: our desires vary in order that our actions will pro duce different results at different times. The role of beliefs is to do with the fact that, given any result, different means are appropriate to that result at different times: our beliefs vary in order that we can choose the most effective me ans at any time to the results that we desire at that time.

 In the end, all selection-based purposes depend on results: to have a purpose is to have been selected by a mechanism which favours certain results. However, the above remarks show that this is true of beliefs only in an indirect sense. For beliefs don't have any results of their own. Rather, their standard purpose is to produce whichever results will satisfy the desires they are acting in concert with.&nb sp; In effect, beliefs get selected at one remove, in virtue of being good at causing actions which cause desired results.

Note that this means that, according to the teleological theory, there is a sense in which the representational pow ers of desires are prior to those of beliefs. Any given desire will be present in order to produce a certain result r, which result is therefore its satisfaction condition. Given this explanation of satisfaction for desires, we can then explai n the purposes of beliefs. Any given belief will be present in order to produce actions which will produce desired results if a certain condition p obtains, which condition is therefore that belief's truth condition.[5]

Let us now return to the example with which I began this section. I have just argued that the biological purpose of any belief is to be present in those circumstances in which the actions it prompts will satisfy desires -- which circumstanc es therefore count as its truth condition. However, the example about not flinching in battle involves a different kind of purpose. For in this case the belief at issue, the belief that you won't be injured, has a extra biological purpose, apa rt from its role in aiding the satisfaction of desires, namely, to ensure that you do not flinch in battle.

In order to deal with such examples, we need to distinguish the "normal" purpose of beliefs, namely, to ensure the satisfaction of desires, from such "special" purposes as stopping you flinching in battle. This distinction then allows us to frame the teleological theory in a way which is consistent with the existence of such special purposes. That is, we should understan d the teleological theory as relating specifically to the normal purpose of beliefs. For, as long as we stick to those normal purposes, then truth is still the requirement for achieving them, in line with the teleological theory of representation.[6]

If you are unconvinced that the belief about invulnerability needs to be true in order to serve its "normal" function, consider the case, say, of Cuthbert Coward. Cuthbert would far rather remain unscratched th an win the battle. Still, if Cuthbert were somehow to be persuaded that he won't be injured (though in fact he will), then even he might be induced to enter the fray. But then he won't get what he desires, which is above all to remain unscathed.&nbs p; It is only the special purpose of getting him to fight, even though he's doesn't really want to, that gets satisfied when the belief is false. By contrast, the normal purpose, of satisfying his desire to remain unscathed, still requires his belie f to be true -- just as the teleological theory, as now proposed, requires.

Of course, Cuthbert has somewhat unsatisfactory desires, from a biological point of view, in the sense that the satisfaction of his desires is unlikel y to further his overall chances of survival and reproduction. This is why beliefs sometimes have special purposes. The point of these special purposes is in effect to by-pass the normal role of beliefs in satisfying desires, and to ensure ins tead that agents with biologically inappropriate desires don't end up performing biologically inappropriate actions. Cowards are a case in point. Their unfortunate desires mean that they are likely to end up running from battle, and thus losin g the any chance of biologically important spoils, just in order to avoid a scratch. And so, in order to protect them against the biological dangers of such consequences, natural selection predisposes them to believe that they are invulnerable, even when the evidence doesn't warrant this belief, so as to stop them performing those actions which would in fact satisfy their desires.

It might seem puzzling that natural selection should give some beliefs two different purposes. Af ter all, natural selection presumably designs biological systems for one ultimate end, namely, the bequest of genes. So why don't beliefs simply have the single purpose of ensuring such gene bequests?

The answer relates once more to the nature of the human decision-making system. Note that this system doesn't work by always choosing that action which is most likely to ensure gene bequests. Rather it chooses that action which is most likely to satisfy existing desires.&nb sp; It is not impossible to imagine biological systems of the former kind, which always aimed directly for gene bequests. But it seems likely that the limitations of our cognitive capacities have prevented us from doing things in this way. Ins tead we aim for such relatively short-term goals as warmth, sex and chocolate ice-cream.

By and large such short-term goals correlate reasonably well with ultimate biological success, which is no doubt why our innate desires, and our ways of acquiring non-innate desires, have evolved as they have. But the satisfaction of our desires won't always coincide with biological success (not all sex leads, or even can lead, to reproduction). And this then means that there are certain b iological risks consequent on our way of doing things. Now, it may be that some of these risks are inevitable by-products of our desire-based decision-making system: for example, it may be inevitable that humans will have extremely strong desi res to avoid injuries, and so inevitable that in certain circumstances this will lead them to act against their biological interests. And this will then lead to natural selection interfering with the normal operation of decision-making system, by gi ving us beliefs which lead us to act in ways that frustrate our desires, but satisfy our biological needs.

Let me sum up the argument of this section. Certain beliefs do indeed have some biological purposes that require them to be f alse. However, this doesn't invalidate the teleological theory of representation. For we can understand the teleological theory as focusing specifically on the normal purposes of beliefs, namely, to guarantee the satisfaction of desires. And these normal purposes don't ever require beliefs to be false.

3.5 The Reality of Beliefs and Desires

In the last section I made a number of definite assumptions about the role of beliefs and desires in our o verall decision-making system. Some readers may want to ask how this tallies with the agnostic attitude to everyday psychology I expressed in chapter 1, when I said that my references to the entities of everyday psychology should be understood merel y as place holders for the true theoretical explanation of human cognition, whatever that may be.

One possible response to this query would be to maintain that the last section's comments about the roles of beliefs and desires need not be read realistically, as committing me to substantial claims about the causal structure of our cognitive system. Daniel Dennett, for instance, argues (1971, 1978, 1987) that everyday psychology commits us only to the "intentional stance", to the view that an individual's behaviour is somehow appropriate to his or her environment and needs, and not to any "design" or "physical" assumptions about the mechanisms that might

be responsible for generating that behaviour. Dennett holds that this inten tional stance is underpinned by general evolutionary considerations, which tell us that our cognitive systems must have some design that will enable us to choose actions that will further our welfare, while leaving open the internal details of that design . On Dennett's conception, then, references to such everyday concepts as belief and desire need not be taken as realistic hypotheses about internal structures, but simply as a way of pointing to the approriateness of actions.

Howeve r, I shall not take this Dennettian line. For one thing, it sits ill with the teleological theory of representation. According to the teleological theory, the representational contents of beliefs and desires depend on how (the abilities to for m) these states have been shaped by natural selection. But if beliefs and desires aren't real states, but only constructs by which we indicate the appropriateness of actions to circumstances, then it is hard to see how natural selection can operate on them. Natural selection favours things which produce certain effects. But it can't favour things which don't exist.[7]

In any case, there is good reason to doubt Dennett's view that everyday psychology is restricte d to the "intentional stance". This relates to a point made in the last section. As we saw, natural selection hasn't arranged our brains so that we always choose actions that are likely to maximize gene bequests. Instead it has fixed on certain relatively short-term goals, like warmth and sex, and on certain ways of acquiring further short-term goals, and arranged for our brains to choose actions which are likely at least to satisfy these goals.

As I observed at the end of the last section, this makes sense from the point of view of natural selection, given that these short-term goals correlate reasonably well with long-term reproductive success, whereas aiming directly for such long-term reproductive success would no do ubt overtax our cognitive capacities (not to mention the cognitive capacities of our evolutionary ancestors). But the fact that the installation of short-term desires constitutes a sensible strategy from the point of view of natural selection should n't obscure the fact that it is a definite design option, a choice of one among a number of different possible internal structures which could ensure that behaviour is more or less appropriate to needs and environment. After all, we can easily enoug h imagine hyper-intelligent non-human beings whom natural selection had made "super-rational", by giving them no short-term desires as such, but simply the sole aim of maximizing gene bequests by always choosing that action which available information ind icated as most likely to achieve that end. And, at the other extreme, we already have terrestrial examples of simple organisms, like insects, with plenty of hard-wired routines driven by short-term needs, but scarcely any ability to modify their beh aviour in response to information about the environment.

So everyday psychology, with its distinction between beliefs and desires, takes us beyond the thought that evolution has somehow arranged that we will choose actions appropriate to our needs and environment, to a specific theory of how evolution has arranged this: evolution has arranged for us to have information about our circumstances, in the form of our current beliefs, and then to choose actions which those beliefs indicat e will satisfy the goals signalled by our current desires. In these respects we are different from the "super-rationalists", since they are not interested in any intermediate goals except gene bequests; and we are different from the insects, i

n that their behaviour is almost entirely insensitive to information about their circumstances.

In the light of these points, I accept that my appeal to beliefs and desires in the last section does indeed take me beyond the stance of chap ter 1, and commit me to the truth of certain basic everyday psychological assumptions as realistic hypotheses. However, now that we have seen why this commitment is inescapable, we can also see why it is unburdensome. For there is plenty of un contentious empirical evidence that everyday psychology is true at just those points where it takes us beyond Dennett's intentional stance. The significant point is not just that everyday psychology says that we are different from the super-rational ists and the insects in having an internal structure of beliefs and desires; in addition, our actual behaviour shows that we are different in this respect. If we didn't differ from the super-rationalists in having desires, then we wouldn't con tinue to act in pursuit of short-term aims, like eating chocolate, even after we know that doing so only makes us fat and so is no help to our reproductive success; and if we didn't differ from the insects in having beliefs, we wouldn't be able to f igure out that one way to acquire some chocolate would be to go to the new confectionery shop around the corner.[8]

This kind of general evidence does not of course confirm every detail of the complex set of assumptions and attitudes which constitute our everyday psychological thinking. But it does seem to me to be enough to justify the kind of core assumptions about the existence of beliefs and desires that I made in the last section. Empirical psychology still has much to discover, both about the more detailed claims made by everyday psychology, and about the "sub-personal" structures by which such everyday psychological claims are implemented. But I don't think it need do anything further to establish t hat human acions are generated by internal causal processes involving beliefs and desires. If our actions were not generated in this way, we would behave quite differently from the way we know we do behave.[9]

3.6 Truth as the Guarantee of Success

In section 3.4 I stressed that the teleological theory of representation needs to be understood as focusing specifically on the role that beliefs play in facilitating the satisfaction of desires, rather than on any further role they may have in fulfilling further biological purposes. However, once we focus on desire satisfaction in this way, then do we still need teleology to explain truth-conditional content? Why not simply explain content directly, by saying it is that property of beliefs which will ensure the satisfaction of desires?

At the beginning of this chapter I argued that functionalism leaves out representation, and that the teleological theory is needed to bring it back in. But perhaps the moral of my remarks in 3.4 about the relation between belief content and desire satisfaction is that we shouldn't start with functionalism in the first place. For what those remarks in effect show is that functionalism presents only a limited picture of the role that mental states play in psychological explanation, a picture which leaves out the role of truth in ensuring the satisfaction of desires. Perhaps once we fill in the missing components of the picture, we won't any longer need teleology to explain representation.

In due course I shall show that this is not so: even after we have paid due accord to the role of truth in ensuring desire satisfaction, we will still need teleology for a full explanation of representation. But it will be worth proceding slowly.

The limitations of functionalism can be brought out by contrasting two different pictures of the structure of action explanation. The first picture, the picture embodied in functionalism, focuses on the internal roles that beliefs and desires play in causing behaviour, and so takes psychological explanation to conform to this pattern:

(A)   X desires G
          X believes that F will produce G
          _____

&n bsp;   X does F.

However, there is also a second picture of the structure of action explanation, a picture embodied in my remarks about the human decision-making system. According to this pict ure, psychological explanation is not solely an internal matter, but also has an "external" structure, which explains, not behaviour, but the achievement of results:

(B)   X desires G
   ;   X believes, of some behaviour, that it will
          produce G
          This belief is true
          _____

          X achi eves G.

If we restrict our attention to "internal" explanations of form (A), as functionalism does, then it scarcely surprising that we become puzzled about the significance of representational notions, since the only role that beliefs an d desires play in (A) is that of causal pushes from the inside, as it were, and not as representers of the external world. But in "external" explanations of form (B), the representational features of beliefs and desires become crucial: the sat isfaction condition of the desire specifies what external result is at issue, the truth condition of the belief specifies how things must be to ensure this result, and the actual truth of the belief specifies that things are indeed so. Far from bein g limited to the internal causes of behaviour, explanations like (B) specify that external circumstances are such as to lead from the agent's behaviour to result G.

This is why we now need to ask whether we really need to appeal to teleol ogy in our theory of representation. The original puzzle that led us to this theory was, in effect, that internal explanations like (A) make no use of representational notions. But now we see that external explanations like (B) do use represen tational notions. And this suggests that we might be able to analyse representational notions purely in terms of the way they enter into such external explanations -- explaining truth, say, as that property which ensures desire satisfaction -- witho ut needing to appeal to teleological considerations after all.

In a moment I shall explain why this doesn't quite work. But let me deal with a minor point first. There is a extensive literature on the question of whether representat ional notions are essential to (B). (See Loar 1981; Devitt, 1984; Field,

1986.) Can't an explanation like (B) always be replaced by a two-stage explanation which first explains behaviour F, as in (A), and then explains G by reference to the fact that F causes G, and thereby omits any explicit mention of truth? Well, maybe so. But the obvious question is why we should want to dispense with truth in this way. The answer, for most of the contributions to the relevant literature, is to do with "deflationary" or "minimalist" theories of truth: defenders of such theories are committed to the replacability of (B)s by (A)s, since they think that mention of truth is always simply a "quotational" variant of what can be said in di squoted terms; while opponents of such theories want to show that (B)s involve ineliminable appeal to truth as a real property of beliefs. My present concerns, however, are orthogonal to this debate. I am not concerned to decide how far talk of truth might be eliminable in favour of something else, but simply to take it at face value, and understand what work it does in our thinking about the world. The question at hand is not whether we can do without truth, but what we do with it . (As it happens, I think that replacing (B)s by (A)s loses sight of a general explanatory pattern, the pattern displayed in schema (B). On the other hand, I don't think that this is the most effective way to argue against the deflationary the ory, given that arguments based on the importance of explanatory patterns are notoriously inconclusive. A better strategy is to press the deflationalist for a theory of translational content. I shall return to this issue in section 3.9 below.)

   The question currently at issue is whether we can analyse representational notions simply on the basis of the way that they enter into external explanations like (B), and without appeal to teleological considerations. Let us consider in more detail how this might work. The idea, in outline, is that truth conditional content might be analysed in terms of the role of truth in ensuring desire satisfaction. We can formulate this suggestion explicitly as follows:

(C)   The truth condition, for any belief, is that
        condition which guarantees that actions based on
      that belief will satisfy the desires it is acting
        in concert with.

   Something like this success-guaranteeing analysis of truth has been p roposed by a number of other writers (Ramsey, 1927, p 29; Putnam, 1978, part 3; Appiah 1986; Mellor 1988; Whyte, 1990.) However, there is an obvious reason why it is not, as it stands, an adequate substitute for the teleological theory. Namely, that it explains truth, for beliefs, only by assuming the notion of satisfaction, for desires. Yet satisfaction is as much a representa tional notion as truth, and so ought itself to be explained by an adequate philosophical theory of representation.

   It is no good simplying offering an account of desire satisfaction parallel to (C), such as:

(D)   The satisf action conditon of a desire is that
        condition which is guaranteed to result from
        actions based on that desire, if the beliefs behind
        the action are tr ue.

For simply adding (D) to (C), without offering any further hold on representational notions, is like trying to solve a single equation with two unknowns. Both (C) and (D) are expressions of the principle:

(E)   Actions based on true beliefs will satisfy the
       desires they are aimed at.

(E) places a mutual constraint on the representational values that a person's beliefs and desires can have. But on its own it does not suffice to pin do wn those values uniquely. If a given attribution of truth and satisfaction conditions satisfies (E), then so will any attribution that simply permutes referents for names and predicates, provided it does so in the same way in both truth and satisfac tion conditions. (Cf Stalnaker, 1984, pp 17-18; Papineau, 1984, p 555.)

   This means is that any theory of representation that explains truth by (C) needs to add something further -- not just (D) -- to explain satisfaction. I add teleology. I explain desire satisfaction in terms of the results that desires are biologically supposed to produce, and then plug this into (C), thus giving truth the biological purpose of satisfying desires.

 There are perhaps other possible options at this point. You might agree with the success-guaranteeing account of truth, as in (C), and agree that something extra is needed, yet disagree that the requisite addition is teleology. However, let us postpone the question of whether (C) can be appropriately supplemented in non-teleological ways until the section after next. For the prior question is whether (C) is even defensible as part of a full account of representation. There a num ber of standard objections to the idea that truth is what guarantees desire satisfaction, which both the teleological theorist and those who want to supplement (C) in other ways need to answer. It will be convenient at this point to deal with these objections.

3.7 Objections to a Success-Guaranteeing Account of Truth

(i) Non-Instrumental Beliefs

Doesn't (C) apply only to beliefs of the form: s will bring about t?   For these are the only beliefs which are directly relevant to the satisfaction of desires, as schema (B) makes clear. Surely, however, an analysis of truth conditions ought to deal with beliefs of all forms, and not just with beliefs about me ans to ends.

   It is not difficult, however, to see why (C) should be considered to hold for beliefs of all forms, as well as for means-ends beliefs. It is true that the relevance of beliefs to actions always depends in the last insta nce on what they imply about appropriate means. And in this sense it is only means-ends beliefs that are directly relevant to actions. But, still, such means-end beliefs, that s will bring about t, will as a rule be inferred by the agent from various other beliefs. And this then institutes the requisite general connection between truth and satisfaction. For if those other beliefs are true, and the inferences from them valid, then the belief that s will bring about t will be true to o, and the resulting action will succeed. So it is a general principle that actions based on true beliefs will succeed, and not just a principle about means-ends beliefs as

such.[10] Consequently, when we invert this principle into an analysis of truth conditions -- analysis (C) -- the analysis promises to apply to beliefs in general, and not just to beliefs of the means-end form.

(ii) Actions Based on More Than One Belief

In general a number of beliefs will lie behind any given action. But this means that the truth of any one belief will be insufficient to guarantee the success of ensuing actions. For desire satisfaction will only be guaranteed if the other beliefs behind the action are also true. So strictly analysis (C) ought to be formulated:

> The truth condition, for any belief, is that
> condition which guarantees that actions bas ed on
> that belief will satisfy the desires it is acting
> in concert with, assuming that any other beliefs it
> is also acting in concert with are true as well.

But this then disqualifies (C) as analysis of truth-conditional representation, for it assumes the notion of truth in explaining it.

It might seem that we could deal with this difficulty by thinking of analysis (C) as applying specifically to cases where single beliefs generate actions on their own, without the assistance of other beliefs (cf Mellor, 1988, p 86). Truth conditions could then be identified as what guarantees satisfaction in such single-belief cases. B ut the trouble with this is that we then run into objection (1) again, since the only kind of beliefs that can generate actions on their own are means-ends beliefs. If we want an analysis of truth that works for beliefs in general, and not just for means-ends beliefs, then we need a way of extending (C) beyond single-belief choices of action.

A better way to deal with the problem is to think of analysis (C) as being applied simultaneously to all the belief types in an agent's repert oire. That is, we should think of (C) as fixing the truth conditions for all those beliefs collectively by, as it were, solving a set of simultaneous equations. The "equations" are the assumptions that the truth condition of each belief will g uarantee desire satisfaction, if other relevant beliefs are true; the overall "solution" is then a collective assignment of truth conditions which satsifies all those equations.

(iii) Can't False Beliefs Satisfy Desi res?

Another initial worry about (C) might be that it makes truth too easy. Surely we don't want to count beliefs as true whenever the actions they prompt have satisfactory results. Can't an action achieve a desired result by acciden t, even though some of the beliefs behind it are false (as when they involve some self-correcting mistake)?

But (C) doesn't in fact rule out this possibility. The suggestion isn't that it's enough, for the truth of a set of token be liefs, that a particular action, prompted by those particular tokens, should satisfy desires. Rather (C) specifies a condition which guarantees, for all tokens of the relevant types, that ensuing actions will satisfy desires.

(iv) Decisions Made Under Uncertainty

In many case an agent will act, not on full beliefs, but on partial beliefs. In such cases the agent's thinking won't pick out any action as certain to satisfy desires, but rather selec t the action that is subjectively most likely to satisfy desires. But then, if the action does succeed, that won't have been guaranteed by the truth of the agent's beliefs about the world.

It is an interesting question as to how far the well-foundedness of decisions made under uncertainty depends on objective features of the world, such as the existence of objective chances. But we can by-pass this issue here. For, once more, there is nothing in (C) which rules out the p ossibility of actions whose success isn't guaranteed by the truth of the beliefs behind them. The idea behind (C) is rather that we should focus on the kind of case where success is so guaranteed, and then analyse truth as what guarantees desire sat sifaction in just those cases. So uncertain decisions issuing from partial beliefs are beside the point. To apply (C) to a given belief, we should stick to cases where that belief is held fully, and figures in decisions which aren't uncertain: truth is what guarantees satisfaction in those cases.[11]

(v) Is Truth Just Pragmatic?

Analysis (C) seems to imply that the virtue of truth is essentially pragmatic, that the reason for wanting tr uth is always so as to satisfy desires. But surely truth can be pursued as an end in itself, and not just because of its pragmatic value. Indeed there are certain questions, about the farther reaches of the universe, say, or the distant past, where our interest in having true beliefs can't possibly be practical, since such beliefs can make no difference to our actions.

But this complaint misses its target. (C) isn't a theory about why we should want truth. It's a t heory of what truth is: namely, for a belief, the obtaining of a condition which guarantees that, if an agent were to act on that belief, the ensuing action would satisfy desires. This doesn't presuppose that anybody will actually act on the b elief. Nor does it presuppose that the only reason for wanting the truth in respect of that belief is to be able to act so as to satsify desires. To be sure, if you do want to satisfy desires, then (C) does immediately imply that you have a mo tive for wanting the beliefs behind it to be true. But that leaves room for other motives for wanting truth, both in the case of practically significant beliefs and practically insignificant ones. In particular, it leaves room for truth to be valued as an end in itself. (Can't we now ask: why should truth be valued as an end in itself? But I take it to be a virtue of (C) that it allows this as a significant question.)

It might still seem that there are some b eliefs that couldn't, even counterfactually, be relevant to an action satisfying a desire. What about the belief that there are no agents, or the belief that all my actions are doomed to failure?[12] At this point we need to appeal to the compositionality of beliefs. As I shall explain in section 3.9, we need to recognize that beliefs are made up of components ("concepts"), the representational significance of which derives from their systematic contribution to the truth conditi ons of the beliefs they enter into, that is, from their systematic contribution to conditions which guarantee that actions based on those beliefs will satisfy desires. Once we recognize this, then we can hope to pin down the representational signifi cance of concepts like agent, doomed to failure, and so on, in terms of their

contribution to beliefs which can be relevant to action, and then use those representational values to build up truth conditions for such special beliefs as can't be relevant to action.

(vi) Non-Natural Beliefs

Analysis (C) applies only to beliefs whose truth is of potential causal relevance to the success of actions. Perhaps this will enable it to accommodate beliefs about the natural world. ; But what about moral, or modal, or mathematical judgements? In what sense, if any, can the truth of such non-natural judgements matter to the success of action?

   I don't propose to pursue this complex topic at this point. Whe ther or not analysis (C) might apply to a given category of judgement depends on the details of the workings of such judgements, and such details are matters of active controversy for moral, modal, and mathematical judgements. I shall offer some fur ther comments on these issues in chapter 6 below.

(viii) Doesn't (C) Presuppose Validity and Hence Truth?

In my answer to objection (i) I appealed to the notion of validity: I argued that analysis (C) could be ex tended from means-ends beliefs to other beliefs because valid inferences from true beliefs of any kind will lead to true conclusions about appropriate means. However, it might be argued that this appeal to validity is illegitimate, on the grounds th at the notion of valdity presupposes the notion of truth.

   Analysis (C) certainly needs the notion of validity. Often agents will draw invalid inferences about means (imagine that they have to decide what to do quickly, or that their situation is very complicated) and then the truth of the beliefs on which those inferences are based won't guarantee the success of their actions. So if analysis (C) is to apply generally, and not just to means-end beliefs, it should strictly be fo rmulated as:

> The truth condition of any belief is that
> condition which guarantees that actions validly
> based on th at belief will satisfy desires.

But this now makes the problem clear: (C) can scarcely be held to constitue an analysis of truth, if it presupposes validity and validity presupposes truth.

   One possible move here might be to deny y that validity does depend on truth. Thus we might seek some purely syntactic notion of validity, defined in terms of some specified structure of rules of inference, rather than the semantic notion of any truth-preserving form of inference. Ho wever, this syntactic strategy seems unpromising. For a start, there are technical difficulties about the completeness of syntactic characterizations of non-first-order validity. And, in any case, given that syntactic characterizations are alw ays answerable to the semantic conception of validity (cf Dummett, 1974), even for first-order validity, it is doubtful that the syntactic strategy will really dispose of the circularity, rather than just brushing it under the carpet.

To deal with this difficulty, I think it is necessary to broaden the focus away from analysis (C) itself, and reintroduce teleological considerations. We need to think of validity as playing a part, alongside truth and desire satisfaction, in fulfillin g the biological purposes of the overall human decision-making system.

It is fairly obvious, on reflection, that this decision-making system needs some mechanism for generating beliefs about means, beliefs that are directly relevant to ac tions, from the total set of background beliefs that may bear indirectly on the achievability of desires. And it will clearly be part of the biological purpose of this mechanism to produce true beliefs about such means, given that the background bel iefs are true. Of course this inferential mechanism won't always succeed in fulfilling this purpose: as I just observed, humans often draw invalid conclusions about which means to adopt. But that doesn't show that validity isn't the infe rential mechanism's purpose, any more than heart failures show that blood circulation isn't the heart's purpose.

As I mentioned earlier (see footnote 5), the biological purposes of beliefs and desires are interdependent, in the sense that desires will only fulfil their biological purposes if beliefs fulfil theirs, and vice versa. We now see that there is a further interdependency, in that both beliefs and desires will only fulfil their biological purposes if the inferential mechanis m fulfils its purpose too, and vice versa. There is of course nothing surprising about such interdependencies. They are a common feature of biological systems. For example, the lungs will only fulfil their biological purpose, of oxygenat ing the blood, if the heart fulfils its purpose, of circulating the blood, and vice versa.

It might not be immediately clear how these observations about biological purposes are supposed to solve the original problem. Don't they jus t amplify the point that truth, in beliefs, and validity, in inferences, presuppose each other, thereby blocking any possiblity of explaining one in terms of the other? But the point of reintroducing biological considerations is not to deny this int erdependence, but rather to show how we can analyse truth and validity simultaneously.

Suppose we start off not presupposing any representational terms like "truth" or "validity". We proceed to describe the workings of the human dec ision-making system. It has various interdependent components: some states (desires) have the biological purpose of prompting actions which will produce specific results; others (beliefs) have the biological purpose of prompting actions which are appropriate to specific circumstances, and hence the biological purpose of co-varying with those circumstances; and then there is an (inferential) mechanism whose purpose is to generate new beliefs out of old ones, under the constraint tha t the circumstances which the latter beliefs are supposed to co-vary with should be guaranteed by the circumstances the former beliefs are supposed to co-vary with. And then having done all that, without using the notions of "truth" and "validity", we can now account for these notions, by saying that beliefs are true when they fulfil their purpose of co-varying with the relevant circumstances, and that inferences are valid when they fulfil their purpose of preserving such truth.

3.8 Alternative Accounts of Desire Satisfaction

This completes my catalogue of standard objections to the success-guaranteeing account of truth-conditional content given by (C). My answer to the last objection

returns us to th e point at which we left the overall argument. For this answer dealt with the difficulty about validity by locating (C) within the biological analysis of the overall human decision-making system. But we have already noted, at the end of the se ction before last, a rather more straightforward reason for making this move. Namely, that (C) on its own simply explains truth, for beliefs, in terms of satisfaction, for desires, and therefore needs supplementation by an independent account of des ire satisfaction. My earlier suggestion was that we should fill this gap too by placing (C) within the biological context of the overall human decision-making system. For this move then allows us to view desires as having a biological purpose, namely to prompt actions which produce specific results, and so enables us to analyse desire satisfaction in terms of this purpose.

A question raised when I made this suggestion was whether this is the only way to remedy the philosophica l incompleteness of (C). Couldn't opponents agree with the rest of my argument, but disagree about the teleology? That is, couldn't they agree that (C) is only part of the truth about truth, which therefore needs to be supplemented with some f urther account of desire satisfaction, but then diverge by offering some different explanation of satisfaction for desires, which does not appeal to considerations of biological purpose?

For example, they might try to identify the results which satisfy desires as those which extinguish those desires (cf Russell, 1921, ch 3; Whyte, 1991). In general, when some desired result is achieved, then that desire disappears. So perhaps we can identify which results are the objects of wh ich desires by reference to which results make those desires go away.

Another alternative would be to appeal to the reinforcement of behaviour (cf Dretske, 1988). Often, when a desire prompts some behaviour which produces a given re sult, that behaviour is reinforced, in the sense that it is more likely to be repeated when that desire next arises. So perhaps we can identifie the results which satisfy desires as those results whose achievement leads to the reinforcement of behav iour.

One problem facing theories of this kind is that they will still face the problem about validity raised at the end of the last section. I dealt with this problem by viewing inferential abilities as part of the overall biologic al system, and accounting for validity in terms of the biological purpose of this ability. Accounts which seek to dispense with considerations of biological purpose obviously cannot offer this solution. Yet they will still face the problem, fo r merely adding an independent account of desire satisfaction to (C) will still leave us with the problem that (C), if it is to work at all, needs implicitly to presuppose an idea of validity, and hence of truth.[13]

There are ot her problems facing the alternative suggestions about desire satisfaction. Take the "extinction theory" of satisfaction first. On the face of it, some desires are only fuelled their own satisfaction (salted peanuts), while others are quenched by their non-satisfaction (sour grapes). Perhaps an extinction theory can somehow be elaborated so as to deal with these prima facie counter-examples. But until this is done, the teleological theory seems to offer a far more powerful and promising approach to desire satisfaction.

As to the "reinforcement theory", it seems odd to view this as a more fundamental account of desire satisfaction than that provided by the teleological theory. The

pheneomenon at issue is that a given action X prompted by a given desire will tend to be repeated just in case that action gives rise to a given result G -- which result the reinforcement theory therefore counts as the desire's satisfaction condition. Now, such reinforcement is certainly a genuine phenomenon. But consider it from a biological point of view. From the biological perspective, reinforcement of some means X amounts to an alternative route to achieving G, alongside the cognitive procedure of noticing that in general X leads to G and acting on this belief. That is, natural selection in effect sometimes arranges for us to acquire a derived desire for X in itself, instead of leaving it to our cognitive system to choose X on the basis o f our prior desire for G and the explicit belief that X is an effective means to G.

This suggests, however, that reinforcement is, in evolutionary terms, a relatively primitive method of generating actions. In section 3.5 above I ha d occasion to observe that we human beings fall short of the kind of biological "super-rationality" which would always choose actions on the basis of explicit beliefs about the most effective way to maximize gene bequests. But at the same time I poi nted out that we have moved some way in this direction, in that we are capable of doing things which we do not desire in themselves, simply because we believe them to be means to things we do desire. To this extent, then, we are more sophisticated t han organisms who rely entirely on reinforcement, and whose only way of benefitting from evidence that X is normally followed by G would be to acquire a derived desire for X.

In view of our greater sophistication in this respect, it would be surprising if our successes in achieving desires were always followed by the reinforcement of the means adopted. Given that we humans can select actions as a result of deliberation as well as conditioning, such automatic reinforcement would be b oth unnecessary and potentially disadvantageous. And in fact it doesn't always happen. Even after much experience of satisfying my desire for chocolate by going to the corner shop, I do not find that I have any desire to vist the corner shop a s an end in itself.

This implies that the reinforcement theory cannot suffice as an account of desire satisfaction. To the extent that some desires can be satisfied without the means adopted being reinforced, as in this last example , we will be unable to equate the satisfaction conditons of those desires with results which lead to the reinforcement of means. If we want a theory of satisfaction that works across the board, we will do better concentrate on those results which desires are suppose to produce when they combine with beliefs in the deliberate choice of actions.

3.9 Do We Need Reified Truth Conditions?

In this section I want to focus on the ontological commitments of the account of representation I have developed so far. (It will be convenient to concentrate on beliefs, but most of the points which follow could be applied to desires too.) We can summarize the account of truth conditional content we have now arrived at a s follows:

(F) The truth condition, for any belief, is that
    condition which guarantees that actions generated
    by that belief will fulfil i ts biological purpose
    of satisfying desires.

Note, however, that this analysis (F) (like (C) before it) refers to "truth conditions", and implicitly views truth itself as a matter of such conditions "obta ining". This creates a prima facie problem. For such talk, if taken at face value, commits us to dubious entities like propositions, or possible states of affairs, or sets of possible worlds.

Some philosophers would be untroub led by commitments to abstract objects like propositions and sets. They can skip ahead to the next section. But I am unhappy with such commitments, for reasons to be given in chapter 6 below, and so in this section I want to try to show that t he reification of truth conditions is not essential to (F).

My argument so far implies that, for any belief-type in an individual's repertoire, an instance of the following schema will hold:

(G) actions generated by that belief will fulfil the belief's purpose of satisfying desires if and only if p

Given this, then one way of understanding analysis (F) is as asserting that claims of the form (G) specify the truth-conditional contents of beliefs. That is, analysis (F) c an be understood as asserting that (G) is an equivalent substitute for:

(H) the belief in question is true if and only if p.

Note now that neither (G) nor (H) refer to truth conditions as such. So if the import of analysis (F) is simpl y that (G) is equivalent to (H), then analysis (F) will be free of any substantial commitment to truth conditions too.

What we want from analysis (F) is a theory of content for beliefs. That is, we want an analysis which explains wh at it is for a belief to have a truth-conditional content, and which therefore gives us a recipe for determining the specific content of any given belief. But we can achieve all this without reifying truth-conditions as objects which attach to belie fs. For we can simply understand (F) as saying that claims like (H), about truth-conditional content, can always be replaced by claims like (G), about biological purposes. I shall understand (F) in this way from now on.

There is a well-known difficulty facing this kind of approach. If we take (G) at face value, and in particular don't read "if and only if" in an inadmissibly intensional way, then we ought to accept such instances as:

th e belief that snow is white will fulfil its
biological purpose if and only if grass is green.

But this is surely unacceptable, if instances of (G) are supposed to amount to specifications of truth-conditional contents.&nbs p; For the truth condition of the belief that snow is white is certainly not that grass is green.[14]

The trouble here, as students of Donald Davidson's theory of meaning will know, is that any "standing belief", such as snow is white, is "always" true, if true at all. So we can get a true instance of (G) simply by mentioning a true standing belief on the left hand side and placing any true statement whatsoever on the right.

This is where we need to recogni ze the compositionality of beliefs. Instead of starting with whole beliefs, and taking analysis (F) to explain truth conditions by equating them directly with instances of (G), we need to start with the components of beliefs, such as singular concep ts, predicate concepts, ways of combining concepts, and so on, and to focus on the referential values of such components, in the sense of the contributions that such components make to the biological purposes of the beliefs they enter into. Analysis (F) can then be viewed as equating truth conditions with conditions built up from such referential contributions. So now we will not construct instances of (G) directly, but only by inference from a set of assumptions about belief components, assum ptions which will specify what is required for the whole beliefs those components enter into to generate successful actions. And then, since we will now be building up the (G)-claim for the belief that snow is white, say, from assumptions about the systematic contribution that the concepts snow and - is white make to success-guaranteeing conditions across the board, we can expect to derive:

the belief that snow is white will fulfil its
biolog ical purpose if and only if snow is white

as desired, rather than:

the belief that snow is white will fulfil its biological purpose if and only if grass is green.[15]

What now of truth itsel f? Those who reify truth conditions as possible states of affairs, or sets of possible worlds, or some such, can simply say that a belief is true just in case its truth condition obtains (the possible state of affairs is actual, the actual world is one of the set of possible worlds, . . .) But those of us who want to avoid reified truth conditions need to proceed more circumspectly. My current thinking on this knotty issue is that we don't need anything more to understand truth itself ap art from an ability to generate the appropriate instance of the schema (H) for any given belief. For, if we are able to do this, then we will have a recipe which tells us what is required for the belief that snow is white to be true, namely, that th is belief is true if and only if snow is white; and what is required for the belief that grass is green to be true, namely, that this belief is true if and only if grass is green; and so on, for beliefs in general. And what more do we ne ed to understand truth, if we have a recipe which tells us what is required for the truth of any given belief?

This is to argue for a version of the redundancy theory of truth, according to which nothing more is needed to understand claim s about the truth of beliefs than to understand that such claims stand or fall with the claims made by the beliefs themselves. It is important, however, to distinguish sharply between the redundancy theory of truth, in this sense, and recent "deflat ionary" theories of truth.[16] The difference is that the redundancy theory leaves room for a substantial theory of content, a substantial theory of what determines the truth conditions of beliefs, whereas advocates of the deflationary theo ry argue that such substantial accounts of content are both unnecessary and misguided.

This difference between the redundancy and deflationary theories is best brought out by focusing on the question of how someone might master the abilit y "to generate appropriate instances of the schema (H)" -- which is how I phrased, at the end of the paragraph before last, the requirement which, according to the redundancy theory of truth, is supposed to render any further understanding of truth

redundant. Deflationalists argue that it is sufficient to know that the sentence +p+ used to identify the belief that p on the left hand side of any instance of (H) should be the same as the sentence +p+ used on the right to specify the requir ement for that belief's truth; or, alternatively, for versions of (H) which specify truth conditions for sentences, that it is sufficient to know that the sentence mentioned on the left of any instance of (H) should be used on the right to specify t hat sentence's truth condition. The redundancy theory, by contrast, is committed to no such "minimalist" account of how to generate instances of (H); it may for instance be combined, as I would combine it, with the view that the appropriate wa y to generate instances of (H) is to accept (H)'s equivalence with (G), and therefore to derive (H)'s instances by determining the biological purposes of the relevant beliefs.

This shows that the redundancy theory should be viewed, not as a competitor to the deflationary theory, but as something on which both deflationalists and their opponents can agree. That is, both sides can agree that nothing more is needed to understand truth itself than a recipe which will tell you for any be lief (or sentence) what is required for its truth. Disagreement arises only on the further issue of what such a recipe need involve. Deflationists think that we need only require that the same phrase appear on the left and right hand sides of (H)-claims. Their opponents will contend that we do not have an adequate recipe for generating (H)-claims until we have a substantial theory of what determines the truth conditions of beliefs (or sentences).

On this issue there seem s to me little doubt that deflationalists are wrong. The point is clearest for the analogue of (H) for sentences. The deflationalist says that you will know how to generate the instances of (H) if you know that the sentence mentioned on the le ft hand side of any instance should be used on the right hand side. But of course this only works if the sentence mentioned is in the language you speak, so that you can use it on the right hand side. To get a notion of truth that applies to s entences in general, and not just sentences of your own language, the deflationalist needs to add that you will get an appropriate instance of (H) if the sentence used on the right hand side translates the sentence mentioned on the left hand side. H owever, this appeal to translation destroys the deflationalist position. For what is it for one sentence to translate another, in the relevant sense, except for them to have the same truth-conditional content? So in order to have an adequate r ecipe for generating (H)-claims, you will need to grasp what it is for two sentences to have the same truth condition. And it is hard to see how you can do this without a substantial account of what determines the truth conditions of sentences.

A similar point applies to the version of (H) for beliefs. The deflationary strategy works fine for beliefs already identified in terms of their truth-conditional contents, as beliefs that p. But for beliefs otherwise identified, in terms of causal relations, say, then we won't know what to put on the right hand side of the relevant instance of (H), unless we have a substantial theory of what determines truth conditions for beliefs in general.

So, while I think th at the redundancy theory gives the right account of truth, I also think that this account needs to be located within a substantial theory of content.[17] The substantial theory of content I favour is in terms of success conditions and biolo gical purposes. However, I don't necessarily want to argue that you need to embrace this specific theory of content to understand the notion of truth. For I

certainly want to leave room for lay people who do not share this philosophical theory of content to understand truth. My view is that such lay people have an "everyday" or "folk" theory of content which is substantial enough to allow a satisfactory recipe for generating instances of (H), but which is philosophically inferior in vari ous respects to the teleological theory of content. I shall not pursue this issue here, however, though I shall return to it in section 3.12 below. For the moment we can simply note that the teleological theory itself is certainly a substantia l theory of content, and so a suitable philosophical setting for the redundancy theory of truth.

One last point about the redundancy theory of truth. As I have explained it, this theory has the disadvantage that it does not provide an explicit analysis of the notion of truth. It tells us that the belief that snow is white is true if and only if snow is white, and the belief that grass is green is true if and only if grass is green, . . . But it does not analyse truth as a property that is common to these and other true beliefs.

If we build up the truth conditions for a given repetoire of beliefs recursively from semantic clauses for the components of those beliefs, then Tarski showed us how to construct a predicate which applies to all and only the truths among those beliefs. This construction, however, makes essential use of set theory. Moreover, it only gives us a predicate equivalent to truth-in-R (where R is the relevant repetoire of beli efs), not a predicate equivalent to truth for beliefs in general. The latter problem can perhaps be solved by equating truth, not with truth-in-any-particular-R, but rather with the second-order property of satisfying-the-correct-Tarski-definition-o f-truth-in-R-for-the-R-you-belong-to.[18] There remains the commitment to set theory. Perhaps there is some way of finessing this problem too. But, rather than digressing further down this by-way, let me simply observe that thos e, like myself, who want to avoid commitment to sets, have the option of abandoning the quest for an analysis of a property common to all true beliefs, and simply settling for what the redundancy theory does undoubtedly give us, namely, knowledge of what is required for any given belief to be true.19

3.10 Broad Contents Revisited

Let me now return to the issue of broad beliefs, beliefs that physical identicals can differ in. As I said at the end of se ction 3.3, the theory of content developed in this chapter will enable us to understand why some beliefs are broad in this sense.

The best way to appreciate the issue of broad beliefs is to return to the contrast I drew between two pictur es of action explanation in section 3.5 above. On the one hand were "internal" explanations, as in:

(A)   1. X desires that G
         2. X believes that F will bring about G
         Therefore,
         3. X does F

If we focus on explanations of this kind, then it is e asy to become puzzled about the existence of broad beliefs. For explanations of form (A) don't require beliefs to do anything except give a causal push to actions from the inside, as it were. And on this conception of beliefs it would indeed b e puzzling that differences outside believers' heads can make any difference to what they believe.

However, as we saw, this isn't the only kind of action explanation. There are also "external" action explanations, which explain, not just means, but results:

(B)   X desires G
        X believes, of some behaviour, that it will
        produce G
        This belief is true
        _____

        X achieves G

Once we focus on thi s kind of explanation, the kind of explanation to which truth-conditional content matters, then the existence of broad beliefs and desires becomes unsurprising. Explanations of form (B) show that truth-conditions are nothing to do with internal push es. Rather, they specify the conditions required for beliefs to satisfy desires. Given this explanatory role, it is easy to understand why some beliefs should have world-dependent contents. For such broad contents will be found whenever two physically identical people are in different contexts in which different conditions are needed to ensure that some piece of behaviour satisfies their desires.

The point is clearest for explicitly indexical beliefs. Su ppose Bill and Ben are physically identical, and that they both have the desire and belief that they express by "I want to be warm" and by "Running around will make me warm". Then they are both likely to start performing the same bodily movements, n amely, running around. And to this extent their beliefs are the same: both beliefs "push from the inside" in the same way. But note now that the conditions that will satisfy their respective desires are different: Bill's desire wil l be satisfied by Bill getting warm, whereas Ben's desire will be satisfied by Ben getting warm. And because of this the condition required for Bill's and Ben's actions to succeed will be different: Bill's action will succeed just in case Bill 's running around will make Bill warm, whereas the success of Ben's action requires the quite different condition than Ben's running around will make Ben warm. And that is why the truth conditions of Bill's and Ben's beliefs are different, despite t heir physical identity. It is simply due to the fact that Bill's and Ben's actions have different success conditions.

This kind of explanation of broadness is not restricted to explicitly indexical beliefs. It will apply whene ver the satisfaction conditions of the desires of two physical identicals are different; for then, as above, the truth conditions of beliefs germane to the satisfaction of those desires will be different too. Given the teleological theory of d esire satisfaction, we can expect this phenomenon to be widespread, even in the absence of explicit indexicality: for the processes which select desires, in genetic evolution and in individual learning, will often select desires because of certain e nvironment-dependent effects of those desires, effects which will not necessarily be present in the different environments of physically identical doppelgangers. So, for example, our desire for water has arguably been selected by a process that favo urs actions that lead us to H2O; by contrast, a being on a planet with XYZ instead of H2O could not have desires which have been selected in this way. Again, it is arguable that my desire for the company of certain people, say, is the result o f learning processes in which those people played an essential role; again, a being who had never met those people could not have developed these desires in this way.

An important special case of broad mental states will be those ac quired in the course of learning a public language. Here we will find mental states whose biological purpose is in essential part to enable us to conform to community usage. (Think of a child being encouraged when it speaks correctly, and disc ouraged when it makes mistakes.) So, for example, I may acquire a concept of arthritis, whose biological purpose is to enable me to apply the word "arthritis" as the rest of my community does. This yields another kind of reason why the mentals tates of physically identicals may have different contents: for somebody may be physically identical to me, and yet live in a community in which "arthritis" is used differently.

3.11 Accidental Replicas

We have just seen how the teleological theory of representation can help us to understand why supervenience is violated by broad beliefs and desires. However, the teleological theory of representation also implies that the supervenience of the mental on brain physics is violated in a far more radical way, a way which is widely regarded as constituting a reductio ad absurdum of the teleological theory.

This more radical violation of supervenience arises because the teleolog ical theory makes representational content depend on selectional history. The content of your beliefs and desires depends, according to the teleological theory, on what purposes they were selected to fulfil. So it follows that another being co uld be physically identical to you, and yet not share your representational states, because it did not share a similarly structured selectional history.

Imagine, to make the issue graphic, that you have a physically identical doppelganger who does not have any selectional history at all, but who simply coagulated out of passing molecules a few moments ago, in some massive cosmic coincidence. Then, according to the teleological theory, this doppelganger will not share any of your con tentful beliefs and desires, despite sharing your physical make-up, since none of its brain states have been produced by any selection processes. And this seems absurd to many philosophers.[20]

An initial point that might b e made on behalf of the teleological theory is that a failure of mind-brain supervenience as such can scarely refute the teleological theory. After all, the example of broad beliefs and desires already shows that the possession of contentful s tates will often require certain kinds of context and history, as well as certain kinds of brain states. So why is it at all surprising that your accidental replica should lack contentful states? Of course, if we still upheld the philosophical view, which was widespread before the recognition of broad beliefs, that differences outside the head cannot matter to mental make-up, then the accidental replica would be a knock-down refutation of the teleological theory. But, as it is, why not s imply accept that the accidental replica is another being whose idiosyncratic background gives it states with different contents to ours?

However, this reply is less than entirely persuasive. The existence of broad beliefs can be de fended on independent grounds, by appeal to pre-theoretical intuitions which owe nothing to the teleological theory of representation. Because of this, the teleological theory is confirmed by its ability to explain of the existence of broad beliefs. However, there are no such pre-theoretical intuitions which show that an accidental replica does not have any contentful states at all;[21] indeed, as I said, most

philosophers view this implication as intuitively absurd. So this implicati on, unlike the existence of broad beliefs, counts against, rather than in favour of, the teleological theory.

   Perhaps defenders of the teleological theory can contest the awkward intuitions about the accidental replica. Intuitions a bout complicated counterfactual situations are notoriously insecure. Can we be sure, when we imagine your accidental replica, that we are really imagining a purely accidental being, rather than one that has somehow been designed, if not by natural s election, then by some supernatural power (such as an omnipotent philosopher who is able to create beings as required to illustrate philosophical points)? If we were imagining such a designed being, then the intuition that it has contentful states w ould be no problem for the teleological theory, for designed states have purposes and so teleological contents. Conversely, if we really are imagining an accidental being, then perhaps we ought therewith to relax the intuition that it has contentful states, which would again let the teleological theory off the hook.

   I shall not develop this line of argument any further, however. For, even if we allow that intuition can somehow simultaneously guarantee both that an imagined bei ng is genuinely accidental and that it has contentful beliefs, there is still a natural way to defend the teleological theory. A defender of this theory can simply point out that the theory is intended as a theoretical reduction of the everyday noti on of represenational content, not as a piece of conceptual analysis. And as such it can be expected to overturn some of the intuitive judgements we are inclined to make on the basis of the everyday notion. Consider, for example, the theoretic al reduction of the everyday notion of a liquid, to the notion of the state of matter in which the molecules cohere but form no long-range order. This is clearly not a conceptual analyis of the everyday concept, since the everyday concept presuppose s nothing about molecular structure. In consequence, this reduction corrects some of the judgements which flow from the everyday concept, such as the judgement that glass is not a liquid.

   This appeal to the idea of a theoretic al reduction might strike some readers as an ad hoc response to the problem of the accidental replica. But this reaction would be unreasonable. For it should have been clear from the start that, if the teleological theory of representation is acceptable at all, it must be as a reduction, not a piece of conceptual analysis. After all, there is clearly nothing about the natural selection of brain states in the everyday notions of beliefs and desires.

   Perhaps the teleologic al theory of representation will one day become part of our everyday concept of representation. By way of analogy, consider the aetiological theory of teleology itself. When, in the nineteenth-century, biologists first started to understand bi ological functions in terms of their Darwinian aetiology, this was inevitably a matter of theoretical reduction, rather than conceptual analysis, since the requisite Darwinian notions were simply not available to pre-Darwinian biological thought. Bu t it is arguable that in the intervening years Darwinian ideas have come to penetrate the concept of function itself, with the result that, to biologists, function now just means: effect for which some trait has been naturally selected. (Cf Ne ander, 1991a.)

   This process, of new theoretical ideas being absorbed into old concepts, is a common enough upshot of the general acceptance of a theoretical reduction. So, as I

said, perhaps one day we will all intuitively think of representation in teleological terms. At which point the accidental replica will cease to be a problem, for our intuitions will then come to tell us that its internal states do indeed lack representational contents (provided, that is, that we succee d in imagining a being who is genuinely accidental). However, all these conceptual changes will happen, if at all, only after the teleological theory of representation has won general acceptance. So for the time being advocates of this theory will do better to rest their case on the arguments for theoretical reduction.

3.12 Empirical Evidence for the Teleological Theory

At this point it might occur to some readers to ask: what exactly is the case for the theoreti cal reduction of representation to teleology? Normally theoretical reductions are supported by empirical evidence. When chemists established that water was H20, for example, they adduced a body of empirical evidence which showed that the exten sions of "water", as used by most people, and "H20", as used by the chemists, were in close agreement. So, by analogy, the teleological theory of representation ought also to be supported by empirical evidence, in particular evidence which shows tha t the teleological theory's ascriptions of content coincide with those made by everyday psychology. But where is this evidence? What grounds have I offered for believing that the everyday desire for r will in fact turn out to have been selecte d to produce r, rather than s, or nothing at all, or the everyday belief that p will turn out to have been selected to be co-present with p, rather than q, or whatever?[22]

In this respect the teleological theory of representation is worse of f than those other theories, discussed in 3.8 above, which agree that truth is the guarantee of desire satisfaction, but then explain desire satisfaction in terms of extinction of desires or reinforcement of behaviour. For, whatever other difficulti es these theories may face, they can at least make a plausible case that they are part of everyday thinking about representation.

The teleological theory of representation, by contrast, needs to be defended as a theoretical reduction, not as a piece of conceptual analysis. So its defenders need to produce empirical evidence that its ascriptions of content coincide with those made by everyday psychology.

I think that they can meet this challenge. But first, before ex plaining the solution, let me say a bit more about the problem.

Defenders of the teleological theory obviously need to recognize that everyday thought embodies a working notion of representational content, which is available prior to any analysis of representation which the teleological theory may offer. After all, everyday thinkers who are quite ignorant of the teleological theory are able to ascribe beliefs, desires, and other contentful states to people, and by and large they are able to agree with each other in such ascriptions.

I take it that such ascriptions are informed by a body of folk psychological assumptions. These will include such general principles as that people act in ways which their beliefs indicate will satisfy their desires; that the truth conditions of belief are conditions which actually produce the satisfaction of desires; that the satisfaction of desires will often, if not always, lead to their extinction, and to the reinfo rcement of the behaviour by which they were achieved; and so on. These general principles

will be supplemented by some more piecemeal truisms, such as that people can normally see what is in front of them, that they normally mean what they say , that they can remember what happened yesterday, that they will desire what they previously desired in similar cicumstances, that they will be thirsty if they have had nothing to drink for days, and so on.

Together this body of everyday knowledge constitutes an implicit grasp of representational notions, a grasp that enables everyday thinkers to ascribe beliefs and desires with specific contents to people. The teleological theory should be understood as offering a deepening and ref inement of this everyday understanding. It deepens everyday understanding, as do all all theoretical reductions, by giving us fuller information about the nature of the reduced phenomenon, information which takes us beyond the surface features by wh ich the phenomenon is normally identified, to the underlying features which explain those manifest appearances. And it refines everyday thinking by adding precision to our assumptions about representation and the propositional ascriptions they infor m.

Let me say a bit more about the way the teleological theory refines everyday thinking. It is an implication of the aguments earlier in this chapter that the general assumptions of everyday psychology do not by themselves yield co mplete determinacy in ascriptions of propositional attitudes. I pointed out in section 3.6 that the assumption that truth guarantees satisfaction places a joint constraint on ascriptions of truth and satisfaction conditions, but that, without some f urther account of desire satisfaction, this constraint can be satisfied by deviant permutations of normal ascriptions of truth and satisfaction conditions. And I argued that the everyday idea that satisfaction extinguishes desires, or the idea that it reinforces the means which achieved them, are not adequate to fill this gap, since they fail to apply to desires in general. In practice everyday thought no doubt fills much of this gap by appeal to such piecemeal rules as that people will desire what they previously desired, that they will be thirsty if they have had nothing to drink for days, and so on. But we can expect that, even so, there will be certain cases where everyday thinking is unable to decide about the content of certain desires, nor, therefore, of the beliefs which inform their pursuit. And in these cases the teleological theory of representation will be able to make determinate what everyday thinking does not. Imagine, for instance, a woman who has a recurring desire which leads her to visit a certain spot in a park. She is not sure why she does this; it could be for the flowers, or the restful atmosphere, or various other reasons. There might be nothing in everyday psychology to determine the conte nt of her desire. But there will still be a fact of the matter as to which previous effect of this desire has led its being preserved, and the teleological theory will fix on this on the content of the desire.

This would be a case w here the teleological theory fills a gap left by everyday thinking. There is also the more extreme possibility that the teleological theory may actively overturn ascriptions of content made by everyday psychology. The accidental replica discus sed in the last section is one example of this. And we can imagine other, more mundane, cases in which everyday psychology's ascriptions of content do not tally with the selectional history of the relevant states, and so are deemed wrong by the tele ological theory. However, to return to the main issue to be addressed in this section, cases like this had better be the exception rather than the rule. For, before the teleological theory can start overturning everyday judgements, we need som e evidence that it is an acceptable theoretical reduction in the first place,

and this requires, as pointed out earlier, reason to suppose that the teleological theory agrees, if not in every case, at least in most of the prior ascriptions of propositiona l content made by everday psychology.

The complaint made at the beginning of this section was that as yet we seem to have no evidence for such agreement. Let me now face up to this challenge. My strategy here will be to appeal to the argument of chapter 2 to provide the requisite evidence. In that chapter I argued that it would be incredible that special-scientific properties should be variably realized, unless their instances are the product of some selection mechanism.   I think that this line of argument will serve to answer our present difficulty. For it implies that it would be incredible that human beings should conform to the assumptions made by everyday psychology, unless their beliefs and desires had b een selected by processes which give them purposes corresponding to their contents.

Before going into details, it is probably worth clarifying the sense in which this argument provides empirical evidence for the coextensionality of the te leological theory's and everyday psychology's ascriptions of content. This relates to the point, originally made in section 2.2, that the "incredibility" of variably realized special-scientific laws without a teleological underpinning is an empirica l matter: the objection to such laws is not just that they offend brute intuition, but, more importantly, that they run counter to the wealth of experience which testifies to the general principle that uniform physical patterns have uniform physical explanations.

To see how these considerations help with the particular problem at hand, recall generalization (E):

(E)   Actions based on true beliefs will satisfy the
        desires they are aimed at.

Now consider an instance involving a desire for some specific physical result, r, like getting hold of an ice-cream:

(I)   Agents who act on true beliefs and the
        desire fo r ice-cream will get some ice-cream.

Note that (I) specifies a uniform physical state in the consequent. Yet the antecedent conditions -- desiring ice-cream, and being a true-believer -- are presumably not themselves uniformly physically realize d. So we might well ask, "Why do these all physically different antecedents have a uniform physical effect?"

This was just the kind of question we asked in chapter 2. And the answer we gave there was that in such cases there w ill always be a selection mechanism which selected the physically disparate instances of the antecedent because they produce the common effect. In fact chapter 2 has already applied this analysis to the specific issue of variably realized desires fo r ice-cream, and argued that the reason the different physical realizations of the desire for ice-cream all lead to the ingestion of ice cream is that this is why they were selected in the first place.

This observation now provides an ans wer to the question of why we should expect the teleological theory to agree with everyday psychology in ascriptions of content, at least in respect of desires. The answer is simply that it would be a mystery that the

desire for some physical result r should do what everyday psychology says it does, as in (I), unless it has been selected to produce r.[23]

The corresponding point about true belief is more interesting. Since generalization (I) generalizes across belief types, not requiring that the agent have any specific beliefs, but just that the agent's beliefs, whatever they are, be true, the "true belief" requirement in the antecedent will be variably realized by the truth of different belief types. Thus, being a true-believer can be realized by: believing that the shop is open and the shop being open; or believing that there is ice-cream in the shop and ice-cream being in the shop; or believing that an ice-cream is within reach and an ice-cream being within reach; and so on. In different cases, different external conditions are required for an agent to be a true-believer, and so for the agent's behaviour to lead to the desired result. And so now we have this version of the variable realizability puzzle: why do all the quite different conditions required for different beliefs to be true all lead, when conjoined with the possession of those beliefs, to the desired result?

And the solution, once more, is that a mechanism has selected those conjunctions of condition and belief precisely because they produce such results. To be more accurate, we should think of the relevant mechanisms as selecting dispositions to form-certain-beliefs-when[24]-certain-circumstances- obtain: for instance, the disposition to form-the-belief-that-an-ice-cream-is-within-reach-when-an-ice-cream-is-within-reach. And the reason why different exercises of these disparate dispositions on different occasions will nevertheless all p roduce the same result, as required by (I), is that these dispositions will have been selected precisely because of the kind of effect the relevant beliefs have when their associated circumstances obtain. It's the conjunction of beliefs and their tr uth conditions that ensures success, and so it's dispositions to form beliefs in conjunction with their truth conditions that is selected.

And this now show us how to answer the challenge of this section in connection with belief contents , analogously to the way we answered it for desires. The teleological theory must match everyday psychology on ascriptions of belief contents because, as before, it would be a mystery that beliefs that p should do what everyday psychology says they do, as in (I), unless they had been selected to be present when condition p obtains.

Perhaps we could have reached this conclusion by a shorter, if less illuminating, route. In the course of this chapter we have had occasion to note that our actions are directed by two kinds of mental states, beliefs and desires: desires have ends attached, and vary over time in ways attuned to our needs, while beliefs tend to "track" specific external conditions; and these beliefs and desires then combine to cause behaviour which causes those ends if those co nditions do obtain. Now, this carefully orchestrated arrangement could scarcely have arisen by chance. If this is really how our psychology works, then surely it must have been designed for that purpose -- not   by a conscious design er, of course, but by the blind selection mechanisms of learning and evolution. (Cf Millikan, 1989a, pp 292-4.) So once more we have prior reason to think that beliefs a nd desires must have been selected for purposes corresponding to their contents, as the teleological theory of representation claims.

3.13 Verificationism Refuted

There is an important general moral to be drawn from the argument o f this chapter, a moral which will be central to the epistemological arguments in the third part of this book. Namely, that the teleological theory is radically at variance with verificationist analyses of meaning which imply a conceptual tie betwee n the truth conditions of judgements and the conditions under which those judgements are asserted. For there is nothing in the teleological theory of representation, when properly understood, to imply that there should be any definite correlation be tween the circumstances in which we are inclined to form beliefs, and those in which those beliefs are true. The reason is that truth-conditional content, for the teleological theory, hinges on the results of beliefs, not their causes. In part icular, the teleological theory identifies truth conditions as those circumstances in which the actions prompted by a belief cause the satisfaction of desires. These are not the same circumstances as those which lead us to adopt the belief. An d there is nothing in the teleological theory to imply any special link between these two sets of circumstances.

   It is true, of course, that there will generally have been some biological pressure in favour of belief-forming processes whi ch tend to yield true beliefs, since true beliefs ensure the satisfaction of desires, and in general the satisfaction of desires is biologically advantageous. But this link is easily disrupted. Most obviously, there is the point that our natur al inclinations to form beliefs will have been fostered by a limited range of environments, with the result that, if we move to new environments, those inclinations may tend systematically to give us false beliefs. To take a simple example, humans a re notoriously inefficent at judging sizes underwater.

Rather more interesting are cases where our systematic tendencies to false belief are themselves the upshot of biological design, rather than simply the result of changed environments . One illustration of this possibility is the belief about immunity to injury discussed in section 3.4. In cases of this kind the normal biological pressure in favour of true beliefs is counterbalanced by a contrary biological pressure, which encourages us to form the belief about immunity even when it is false, so as to get us to fight and win. And there are many other similar[25] cases in which biological pressures produce systematic inclinations towards false beliefs. These fa lse beliefs then lead us to act in ways that frustrate our desires, but tend to further our biological needs. And in consequence the circumstances in which we form such beliefs will be systematically different from those which make them true, for tr uth conditions are tied to the satisfaction of desires, rather than biological needs.

   Verificationists might feel inclined to respond that these observations are beside the point, on the grounds that verificationism only asserts a tie bet ween truth and normative assertion conditions, not actual ones. What matters is when people ought to assert claims, not when they do. So biological demonstrations that people often do assert false claims are beside the point. (After all, verificationists can point out, a distinction between "canonical" assertion conditions and actual practice has always been implicit in verificationist thinking, for without some such distinction verificationism will fail to leave any room for false judge ments.)

However, I don't think that an appeal to this kind of distinction can save verificationism from the biological facts. For once we allow the kind of radical gap between truth and assertion that is implied by biology, then ver ificationists face the problem of providing some independent grounding for assertoric norms. It is one thing to allow, say, that individual assertoric practice sometimes falls out of step with the majority line. For then the majority will prov ide the norm for individual

practice. But if verificationism accepts that there can be judgements which nearly everybody gets wrong nearly all the time, then what basis is left for the thought that nevertheless there are agreed standards of correct judgement which are conceptually tied to the truth?[26]

Of course, there is one way of construing "assertoric norms" which will create a conceptual link between truth and normative conformity -- namely, we can equate such norms with whi chever judgemental procedures will lead us to the truth, and then giving some independent analysis of truth. This is how I myself think of assertoric norms, with the independent analysis of truth being provided by the teleological theory of represen tation.

But this is not verificationism. Verificationism aims to proceed in the opposite direction, by given some self-standing account of assertoric norms, and then defining truth in terms of the satisfaction of such norms. T he normal basis for such a verificationist account of norms is the actual assertoric practice of the community. My point is that this route ceases to be available once verificationism concedes to the teleological theory that the whole community can usually be wrong.

1. I am interested in representation as a problem for physicalism. It is worth observing, however, that the problem is scarcely peculiar to physicalism. Even dualists, for example, have an obligation to e xplain how their special mind-stuff can stand for other things. Not that they have always recognized this problem, no doubt because their mind-stuff had so many special powers anyway -- such as the ability to exist outside space but in t ime, to be transparent to itself, and so on -- that one more special power scarcely seemed worth worrying about.

2. Cf Fodor (1990, pp 63 ff).

3. Versions of this teleological approach to mental representation are found in D ennett (1969, ch 9; 1987, ch 8), Fodor (1984), Millikan (1984, 1986, 1989a), Papineau (1984, 1986b, 1987), McGinn (1989, ch 2). Fodor has since recanted. He now holds (1990, Ch 3) that the teleological approach fails to solve the disjunction p roblem. He says that there is nothing in teleology to tell us that a frog's fly detector, say, represents flies rather than flies-or-any-other-small-black-moving-objects, since a properly working detector will respond to any small black moving thing . But Fodor is here assuming that the purpose of the fly detector is fixed by what causes it, rather than by what it is supposed to cause. However, as I shall stress in what follows, biological purposes are always a matter of results. In particular, the purposes of beliefs are to get the organism to behave in a way appropriate to certain circumstances. This is why the frog's detector registers flies: the frog's states cause the frog to behave in a way appropriate to flies, an d not just to any small black inedible dots. (Why flies, rather than food, or survival, or gene perpetuation? This is a different question, about a "vertical" indeterminacy which is orthogonal to the "horizontal" indeterminacy of the disjuncti on problem. I shall answer it in footnote 8.)

4. This arguably oversimplifies the example somewhat (cf footnote 25 below). But for the time being it will be helpful to sacrifice biological realism to explanatory convenience.

5. While t his gives us one sense in which desire satisfaction is prior, there are other senses in which the representational powers of beliefs and desires are mutually dependent. The sense in which desire satisfaction comes first is this: the biological aim of desires is not (except in special cases) to produce true beliefs, but the biological aim of beliefs is standardly to satisy desires. However, this is consistent with the point that desires always act in concert with beliefs when prompting ac tions, just as much as vice versa, and therefore that any desire fulfilling its biological purpose will depend on beliefs fulfilling their biological purposes too. Moreover, because of this, we can expect desires and beliefs also to be psychodevelop mentally interdependent, each category becoming differentiated as a distinct psychological state only when the other is.

6. Ruth Millikan (1984, 1989b) uses the phrase "proper function" for those effects of biological traits which they have been selec ted to produce (that is, for the aetiological notion of "function" or "purpose"). It is perhaps worth observing that in this sense both the "normal" and "special" purposes of beliefs are "proper functions".

7. Dennett himself favours a selection ist account of representational content (1969, 1987). However, he seems not to have noticed the tension between this and his non-realism about beliefs and desires.

8. These points about our human structure of beliefs and desires now answers the question raised at the end of footnote 2, and explains why our different desires have different satisfaction conditions, rather than all being aimed alike at the ultimate evolutionary end of gene perpetuation. For while it is true that the biologica l purpose of all desires is in the end to foster gene bequests, different desires have been designed to foster this end in determinately different ways. This shows up in the fact that the desire for chocolate, say, doesn't disappear when you accept that eating more chocolate won't help you pass on your genes. The appropriate way to think of the purpose peculiar to a given desire is as that result the desire will lead us to pursue whether or not we believe that result is a means to futher ends. For further discussion of this point, see Papineau (1987, sect 4.3).

9. Dennett's non-realism about belief-desire psychology makes him think that it is absurd to suppose that empirical discoveries might show that we don't have beliefs and desir es (1987, p 233-235). I agree that this is absurd. But this is not because I agree with Dennett that belief-desire psychology is non-theoretical and so somehow insulated from empirical evidence. Rather, I think it is theoretical, but alr eady established by a wealth of evidence.

10. I owe the argument for this principle to Horwich (1990).

11. This response to the objection about uncertainty was suggested to me by Hugh Mellor.

12. These examples were put to me by David Owens an d David Sanford respectively.

13. Whyte (1990) aims to deal with this problem by arguing that the causal roles by which we ordinarily identify beliefs, and which then fix their success conditions, happen specifically to involve valid rather than inval id inferential moves. I agree that common sense psychology regards the valid implications of beliefs as constitutive of those beliefs, by contrast with any characteristic tendencies to generate invalid

conclusions. But I think this is because ordinary thought identifies beliefs by their truth conditional contents, and then helps itself to the idea of those consequences which validly follow. This means that any attempt to reduce truth conditional content cannot appeal to common sense psyc hology's view that certain inferential consequences are constitutive of the identity of beliefs. For these sets of constitutive consequences cannot be characterized without the notion of semantic validity.

14. Why shouldn't we read the "if and only if" in an intensional way? Well, if we read (G) as saying that ". . . the belief . . . will fulfil its biological purpose in all possible worlds where p", this will solve the snow is white/grass is green difficulty, but only at the cost of introducing possible worlds. It is true that an explicit reference to possible worlds is only one possible way of analysing ". . . the belief . . . will necessarily fulfil its biological purpose if and only if p". In chapter 6, however, I sha ll argue that, whether or not we adopt the possible worlds analysis of modality, modal judgements cannot be viewed as legitimate expressions of belief, and so are ineligible for essential roles in our best theories. So I prefer to solve the snow is white/grass is green difficulty without using modal notions.

15. Donald Davidson's approach to meaning (1984) can also be viewed as offering a kind of analysis of truth-conditional claims like (H) (rephrased to apply to sentences rather than bel iefs), through not a reductive analysis, as above, but rather an implicit analysis, via an explanation of how to test an empirical "meaning-theory" which specifies (H)-claims for all a community's sentences. (For an exposition of this interpretation of Da vidson, see Papineau, 1987, sections 2.4-8). This Davidsonian approach to truth-conditional content has extra difficulties with the "snow is white/grass is green" problem, however. For, while the problem can still be solved, given strong enough requirements about the need for "meaning-theories" to derive their (H)-claims from separate assumptions for sub-sentential components, it is unclear how to motivate these requirements within the Davidsonian programme. From my perspective, th is is not a difficulty. I take a realistic view of belief components and their referential contributions to biological purposes. So I don't need any independent justification of the compositionality requirement, of the kind essayed by Davidsonian theorist s, such as that native speakers, or perhaps meaning theorists, need to derive their knowledge of an indefinite number of (H)-claims from a finite amount of sub-sentential information. From my point of view such doubtful appeals to the preconditions for knowledge of meaning-theories are irrelevant, since I take the compositionality requirement to be a direct upshot of the semantic facts, irrespective of whether or not any native speakers, or meaning theorists, know a theory of those facts.

16. Fo r the redundancy theory, see Ramsey (1927). Deflationary theories are defended in Quine (1970), Leeds (1978), Horwich (1982, 1990); for a general discussion, see Field (1986).

17. From this perspective, the redundancy theory can also be vi ewed as consonant with the idea that truth involves correspondence with the facts. Those who adopt the redundancy approach to truth will not, of course, want to explain truth in terms of possible facts "obtaining". But it seems natural to say, given the redundancy theory, that when the belief that snow is white is true, for instance, this is in virtue of the fact that snow is white. This doesn't explain truth in terms of facts, but rather introduces facts as what make true beliefs true;& nbsp; still, when a belief is true, there will be a

corresponding fact. There remain questions about the "thickness" of any such fact; are there any other reasons, apart from the truth of the corresponding judgement, for recognizing the fact, such as, say, its causal significance?

18. Since this definition generalizes over Rs, the "adequacy condition" which provides an "external" test for the correctness of Tarski-style definitions of truth-in-particular-Rs will become part of the definiti on of truth-in-general. This is what we should expect: it reflects the point that you cannot have a general recipe for generating (H)-claims without a substantial theory of what determines the contents of beliefs.

19. At one time I thought that the we could equate truth for beliefs in general with the property of "generating actions which are guaranteed to succeed" (cf 1990, p 30). But I now think that the "guaranteed" here conceals a reference to a reified truth condition, since wha t we need, for the truth of any token of a belief type, is not just some actual fact that will cause the action based on the belief to satisfy desires, but, more specifically, the obtaining of that possible fact which guarantees success for all actions ge nerated by tokens of the belief type. An analogous point would apply if we tried to equate truth for beliefs with the property of "fulfilling their biological purposes of satisfying desires"; for such fulfilment needs to be understood in terms of the general condition required for the relevant belief type to fulfil its purpose, not just in terms of any accidental route to desire satisfaction.

20. Cf Cummins (1989, ch 7); Whyte (1993). The problem of the accidental replica is also dis cussed by Millikan (1984, p 94). Note that the accidental replica wouldn't present a problem if we divorced the teleological theory of representation from the aetiological theory of teleology. But since I see no virtue in non-aetiological acco unts of teleology, I shall not pursue this option further.

21.This is perhaps a bit strong, given that some of the arguments for broad mental states do arguably have the corollary that your accidental replica will lack some of your mental states. ; Thus, if the broadness of your concept of water depends upon which liquid was around when you learnt this concept, then a being that never learnt anything couldn't share your concept. Still, many other broad attitudes don't depend on learning in t his way, and so there will be no pre-theoretical reason to deny them to your accidental replica. And, apart from that, plenty of your beliefs and desires aren't broad at all, and so intuitions about broadness will do nothing to explain why your repl ica lacks these.

22. I owe this objection to the teleological theory to a conversation with Andrew Woodfield.

23. So far this only deal with desires for physical things. But the story can be elaborated to accommodate desires for non-physical things, provided those non-physical things in turn have physical effects, by reference to which the desires in question can then be selected. A similar point applies to beliefs. Beliefs can be selected to be co-present with non-physical condi tions, provided those conditions have physical effects by reference to which such co-presence can be selected. The discussion of hierarchies of selection mechanisms in section 2.8 is relevant here.

24. "When" only makes immediate sense for index ical beliefs. For standing beliefs, we need the compositionality of beliefs to give it substance. That is, we need to remember that beliefs are made of components, whose representational significance

depends on their systematic contribution to the truth conditions of those beliefs, and that what gets selected, in the first instance, are therefore dispositions to deploy such components in just those cases when their contribution to the truth condition of the resulting belief will be satsified.

25. In fact the earlier description of the immunity-from-injury case was something of an oversimplification. The real biological problem in such cases is not that there is no psychological desire corresponding to the relevant biological need (to fight and triumph), but rather that this desire is insufficiently strong in comparison with other conflicting desires (like wanting to avoid injury). Our biology then compensates by favouring beliefs that will get us to pursue such insufficiently st rong desires, even on scanty evidence. (Perhaps the best-known of the many other examples of this structure is the human readiness to conclude that given foods are poisonous, thereby compensating for our biologically inappropriate tendency to let ou r hunger outweigh our fear of poisoning.)

26. Followers of Michael Dummett might feel inclined to argue that there must be such standards, in order for people to be able to acquire or manifest their grasp of judgements. (Cf Dummett, 1976, p 101.) But the thesis that acquisition and manifestation depend on agreed standards of correct judgement is itself undermined by the observation that there are judgements which everybody tends to get wrong.