Naturalism and the Mental

MICHAEL TYE

One project which many contemporary philosophers take to be of cardinal importance is the development of a satisfactory naturalistic theory of the mind. Without such a theory, it is feared, the mental will remain forever enigmatic—or, more radically, if the natural world is taken to be all that there is, the deeply ingrained conception we have of ourselves as undergoing and acting on mental states will be threatened.

Notwithstanding the widespread agreement that a naturalistic theory of the mind is needed, there is considerable disagreement about precisely what sorts of items in nature mental states are most plausibly taken to be. One straightforward proposal for naturalizing the mental is that it be identified, at both token and type levels, with the physical. However, the suggestion that mental state types are identical with physical state types is generally regarded as untenable. Instead the most popular view today is that mental states are functional states. It is primarily over how this view is to be spelled out further and how the intentional aspects of mental states are to be naturalistically captured that debate of late has centered.

My aim in this paper is to undermine the currently popular philosophical project of naturalizing the mental via some sort of type reduction or analysis. This is not because I think that the mind is some eldritch, nonnatural entity whose workings are beyond our understanding. On the contrary, it seems to me evident that the mind has a place in the natural world. Where many philosophers have gone wrong, I maintain, is in believing that psychological naturalism is defensible only if it can be shown either that mental state types have non-mental scientific essences or that our psychological concepts are susceptible to exhaustive, or partial, reductive philosophical analysis.

The belief that psychological naturalism requires appropriate type level reductions or analyses is, I might add, fast becoming a dogma in the philosophy of mind. And whether or not it is true that we should let sleeping dogmas lie, this is, I suggest, one dogma that is very much in need of disturbing.

The structure of the paper is as follows: I begin by laying out what I take to be the three major brands of fully reductive philosophical naturalism with respect to the mental. After explaining why I am opposed to these views, I turn to a fourth brand of naturalism that is conceptually driven but only partially reductive. This view is also one that I reject. I then go on to develop and defend a naturalistic approach to the mental which is in keeping with our ordinary, pre-theoretical conception of naturalism and which has the effect of rendering all four of the repu-
What is a naturalistic theory of mental states? Let us begin with a look at what Ned Block (1980b) says are the concerns of the reductionist doctrines of behaviourism, functionalism, and physicalism:

All three doctrines address themselves to such questions as “What is pain?”—or at least to “What is there common to all pains in virtue of which they are pains?”

It is important to note that metaphysical functionalism is concerned (in the first instance) with mental state types, not tokens—with pain, for instance, and not with particular pains…Most functionalists are willing to allow that each particular pain is a physical state or event…Where functionalists differ from physicalists, however, is with respect to the question of what is common to all pains in virtue of which they are pains. The functionalist says the something in common is functional, while the physicalist says it is physical (and the behaviourist says it is behavioural). (p. 172)

It is not entirely clear from these comments whether Block intends the universal quantifier in the question “What is common to all pains in virtue of which they are pains?” to range over actual pains or actual and possible ones. The counter-examples that Block later considers, however, to various different versions of functionalism (counter-examples involving imaginary snake-like creatures that communicate by emitting strong magnetic fields, homunculi-headed robots, and Martians) make it clear that he takes functionalism to yield definitions of mental states (see his 1980a). So, the question that functionalism addresses, as Block understands it, must concern both actual and possible pains.

Now functionalism, behaviourism, and physicalism are normally regarded as naturalistic theories. So, generalizing from Block’s comments, we might propose that naturalistic theories of mental states are attempts to provide reductive answers to questions of the sort “What is X?” or “What is common to all (actual and possible) X’s in virtue of which they are X’s?”, where X is any mental state. A naturalistic theory need not be a physicalist theory, however, on the above conception. What is required rather is that the vocabulary in which the answers are framed be entirely non-mental (for example, behavioural or functional).

The root assumption of naturalism, then, on the above view, is that mental states have essences that are expressible in non-mental language. The philosophical project of naturalizing the mental is the project of specifying these essences, of stating necessary and sufficient conditions for the correct application of mental state terms.
This conception of naturalism is intended to cover not just non-intentional mental states such as pain but also intentional mental states such as the state of believing that snow is white. For states of the latter sort, the project is to naturalize both their species and their representational aspect. So, for example, among the questions to be addressed are these: “What is common to all beliefs in virtue of which they are beliefs?” and “What is common to all beliefs that represent that snow is white in virtue of which they represent that snow is white?” Jerry Fodor has claimed that there is some consensus about what it would be like to have satisfactory answers to questions of representation. He says:

The worry about representation is above all that the semantic (and/or the intentional) will prove permanently recalcitrant to integration in the natural order...What is required to relieve the worry is therefore, at a minimum, the framing of naturalistic conditions for representation. That is, what we want at a minimum is something of the form “R represents S” is true iff C where the vocabulary in which condition C is couched contains neither intentional nor semantic expressions. (1990d, p. 32)

So, to naturalize intentionality, we must provide reductionist necessary and sufficient conditions.

How are different proposals to be evaluated? If we suppose that mental states have essences that may be discovered by a priori reflection then the test of a proposal is whether it accords with what competent users of mental state terms would say about their applicability in a variety of actual and counterfactual cases. For a priori discoverable essences are ones that match our concepts of mental states, and these concepts determine our categorizations of cases as involving tokens of the mental states or not. This seems to be Fodor's view, for the case of representational aspects of mental states, in the passage quoted above. On the other hand, if we suppose that mental states have essences that cannot be completely described without scientific investigation then the situation is more complicated. For the philosophical project of naturalizing the mental can no longer depend on conceptual analysis alone. Instead it must be supposed that the project is to delineate a priori what sorts of non-mental essences mental state types have (for example, functional or physical) and to indicate which sciences will tell the rest of the story (for example, cognitive psychology or neurophysiology). Naturalists who take this line typically hold that, under any acceptable account of the mental, all or, at any rate, most of our ordinary everyday speech dispositions with respect to the use of mental language in both actual and counterfactual situations must be preserved. So, thought experiments about whether or not given creatures undergo given mental states are still very much relevant to the assessment of individual proposals.

We might call the first of the above two views “Analytic Naturalism” and the second, in its usual form, “Conceptually Regulated Scientific Naturalism”. In the

1 Later in the paper Fodor adopts, albeit tentatively, a version of the causal covariation account of representation.
former case, direct examination of our mental concepts, and the necessary and sufficient conditions that govern their application, is taken to yield a priori reductive analyses of the essences of mental states. In the latter, scientific investigation, together with philosophical reflection regulated by our pretheoretical conception of mental states, is needed to come to a full understanding of their essences.

There is a second variant of scientific naturalism that deserves mention. One might hold that our ordinary conception of mental states need not be respected (even partially) by any final theory of their nature. On such a view, there is not much for the philosopher to do. The task of the philosopher is simply to speculate in a very general way as to what sorts of non-mental essences mental states have, but to leave the rest of the story to science. If the combination of philosophical speculation and scientific investigation ultimately results in an account of the mental that is badly at odds with our ordinary conception, so much the worse for the ordinary view.

Let us call this position “Conceptually Indifferent Scientific Naturalism”. The position is a reductionist one, since it takes mental states to have hidden essences that are describable in non-mental language, but it eschews any conceptual analysis or common sense conceptual regulation.

So much, then, for variants upon the general project of revealing the essence of mental states in naturalistic terms. The three views I have now distinguished are the major versions of fully reductive philosophical naturalism with respect to the mental. All of these views are, I believe, problematic in one way or another; furthermore, none of them are needed in order to secure a place in nature for the mental. It is to a defence of the former claim that I now turn.

II

I noted in the last section that Analytic Naturalism is committed to the thesis that our psychological concepts have necessary and sufficient conditions for their application—conditions that may be elicited by a priori examination. This thesis does not sit well with the long and miserable failure of philosophical reflection to have produced any unobvious, satisfactory conceptual analyses. Surely the lesson to be learnt from the past here is that concepts generally don’t have a priori discoverable necessary and sufficient conditions. This is also strongly suggested

2 Such a view has been held by many philosophers. See e.g. Carnap (1932), Armstrong (1968), Lewis (1972). It is also held for the intentional aspects of mental states by, e.g., Dretske (1988) and Stalnaker (1984).


4 A position of this sort is held by William Lycan, for example. Lycan (1988, p. 32) claims that ordinary mental words like “belief” and “desire” pick out real states whose nature will be revealed by psychology, but states which may well turn out to lack most of the properties attached to them by commonsense.
by recent work in cognitive psychology on categorization and concept possession by Eleanor Rosch and others (see Rosch 1973, 1975, 1978; Rips 1975; Smith and Medin 1981).

So, unless it can be shown that there is something very special about our ordinary concepts of the intentional and/or non-intentional aspects of mental states, it is, I suggest, quite implausible to suppose that they have definitions which may be revealed by armchair investigation. Instead, one worthwhile empirical hypothesis deriving from Rosch is that everyday psychological concepts are prototype concepts, that is, concepts whose application in given instances is governed by tacit judgements about whether there is sufficient similarity between those instances and prototypical or paradigmatic cases.\(^5\)

Now it might be argued that there are some analytic naturalists who are not subject to the above criticism. I have in mind here those philosophers who may be classified as a priori commonsense functionalists with respect to the various \textit{species} of mental state (worry, pain, belief, desire, etc.), and who are prepared to grant that what makes something a token of a given mental state is simply its having \textit{enough} of the relevant causal features. Such an approach is tantamount to the thesis that mental state terms express prototype concepts, where the prototypes draw on certain sorts of common sense causal features and no others.

Now it is certainly not \textit{obvious} that ordinary psychological concepts are prototype concepts with the above character. Indeed it seems to me that there is good reason to deny that this is the case. Let me explain.

It is questionable whether creatures that are neurophysiologically very like us but that inhabit immediate environments very unlike ours (for example, brains in vats) can be subject to beliefs and desires with our representational contents.\(^6\) But it seems reasonable to suppose that they can have some beliefs and desires; and they can also be subject to non-representational mental states, for example, pain.

Consider, then, the case of an envatted brain. According to the commonsense functionalist, as we are now understanding her, the brain itself experiences pain if, and only if, it is in a state that is sufficiently similar to prototypical instances of pain with respect to its common sense causal role (where this role may ultimately be specified without the use of any psychological language). But the brain’s internal states are not brought about by bodily damage and they do not issue in any motor behaviour. Moreover, the functionalist cannot appeal to other psychological states causally connected to pain (for example, the desire that pain cease). For the brain lacks, in turn, the appropriate commonsense inputs and outputs for these psychological states. The envatted brain, then, is never in a state that is in any way similar to prototypical instances of pain with respect to their non-mental common sense causal features. So the counter-intuitive result is

---

\(^5\) On Rosch’s view, none of the features of prototypical \textit{F}’s are either necessary or sufficient conditions for being an \textit{F}. As Rosch acknowledges, this position finds its roots in the writings of the later Wittgenstein. There are, I might add, other views besides the Roschian one that are worth considering. For alternative empirical hypotheses about concepts and categorization, see Smith and Medin (1981).

\(^6\) See here my comments on p. 437 and in footnote 15.
reached that the brain is never in pain, even if it is neurophysiologically very like human brains and is, from time to time, in the brain state or one of the brain states normal humans are in when they are in pain. And what goes for pain goes mutatis mutandis for other species of mental state.

The common sense functionalist is not without responses to the specific difficulty just raised. It illustrates one dimension of a much more general problem, however—a problem that has not, I think, been satisfactorily solved by the commonsense functionalist. Even if the view is understood in the rather loose prototype way I am suggesting, it is still intended to yield conditions (however qualified and vague) that are necessary and sufficient for the application of mental state terms. But it seems intuitively wrong-headed to insist that all possible creatures that have mentality (including brains in vats, angels, gods, extra-terrestrial with the physical appearance of clouds) must share the ordinary, everyday sensory inputs and behavioural outputs that are associated with actual human beings. Indeed, it seems a piece of a priori dogma to assert that filling enough commonsense causal/functional roles is strictly necessary for having mental states even if it is sufficient.

Let us now prescind from Analytic Naturalism and turn instead to Conceptually Regulated Scientific Naturalism. According to the latter view, mental state types have non-mental essences. The task of the philosopher of mind is to specify what sorts of essences these are and correlative to say which sciences will discover them, the primary constraint on any acceptable proposal being that it must be compatible with our ordinary, pretheoretical views about where the boundaries of the mental lie.

The major problem with Conceptually Regulated Scientific Naturalism is simply this: If our commonsense categorizations of actual and possible cases are to be respected, there can be no plausible scientific reduction of mental state types in non-mental terms. For where would such a reduction be found? If we suppose that mental states have neurophysiological essences then neurophysiology will tell the story. But, as Putnam emphasized some time ago (see his 1980), creatures come from such diverse species (humans, bats, dolphins, octopuses) that it is an unwarranted leap of faith to suppose that even in the case of a simple state such as pain—a state we ordinarily think of many different sorts of creatures as having the capacity to undergo—there is a single common neural state. And, of course, once we consider possible creatures, the situation becomes absolutely untenable.

So, mental state types do not have neurophysiological essences. And, for parallel reasons, they do not have essences at the levels of physics and chemistry either. What about at the level of cognitive psychology? Might mental state types have functional essences that are extractable from the theories of cognitive psychology? The answer again is surely “No”, if our ordinary intuitions about mentality are correct. Cognitive psychologists develop theories about the exercise of our cognitive capacities, about how we understand language, how we generate mental images, how we see, and so on. The experiments they conduct are per-

---

7 For one possible response, see Tye (1983).
formed not surprisingly on us, that is, on fellow human beings. There is no a priori reason why these experiments should have the same results for creatures from other actual species on which they could be performed. And however the experiments come out in actual cases, we can surely imagine them coming out differently for other possible creatures. That is, we can surely imagine creatures whose underlying cognitive processes and mechanisms are very different from ours, but who nonetheless are subject to everyday psychological states.

Another familiar difficulty facing attempts to define mental state types in "psycho-functional" terms concerns the general requirement that, for each mental state type, there be certain non-mental inputs and outputs that flesh out the functional definitions. For again it seems implausible to maintain that there are any inputs and outputs that are common to all actual and possible cases of pain, belief, fear, etc., even if non-commonsense inputs and outputs are now permitted. So, whichever science is chosen to specify the non-mental essences of mental states, there are grave difficulties with the position of Conceptually Regulated Scientific Naturalism.

That brings us to the next naturalistic alternative. According to the view I earlier labelled "Conceptually Indifferent Scientific Naturalism", mental states may well turn out not to have most of the properties we ordinarily attribute to them. Moreover, even if they don't turn out this way, it is certainly not metaphysically or conceptually necessary that they have such properties; and neither is it sufficient. So, any conclusions we draw from thought experiments which rest on intuitions we might have about the mental states of non-actual creatures in the light only of our ordinary, everyday conception of those states may well be in error. In matters of the mental, science, together with philosophical theorizing based on it, can be our only guides.

There are several serious problems with this view. To begin with, Conceptually Indifferent Scientific Naturalism assumes that mental state types have hidden non-mental essences. But how are we to decide what sorts of essences these are? If they are physico-chemical then some of our actual attributions of mental states to creatures belonging to other species are almost certainly wrong. This is a consequence that the Conceptually Indifferent Scientific Naturalist is prepared to swallow. After all, as I noted earlier, defenders of the view are prepared to assert that our commonsense conception of belief and desire may be badly in error. Why not then say that mental states have physico-chemical essences? But equally why not say that they have psycho-functional essences? Once we remove all conceptual constraints, it appears that just about any answer will do. And this, it seems to me, is the mark of a theory that is out of control—a blind attempt to save naturalism whatever the cost.

Secondly, it is a direct consequence of the present version of naturalism that our ordinary, pre-theoretical views about what is going on in familiar thought experiments involving the mental must be mistaken. Consider, for example, Block's well known thought experiment (in 1980a) concerning friendly Martians who land on earth and whom we come to interact with in much the way in which
we interact with one another. It seems very natural to suppose that the Martians are subject to emotions, beliefs, and desires, even if it turns out that there are very significant underlying physiological and functional differences between them and us. But, according to Conceptually Indifferent Scientific Naturalism, this supposition cannot be correct. An immediate tension thus arises between this form of naturalism and ordinary reflection.

Now there is no doubt that the Conceptually Indifferent Scientific Naturalist would simply reject the deliverances of ordinary reflection in thought experiments of the above sort. But this rejection does place a heavy burden of proof on such a philosopher. Why, on earth, should we accept a view that goes so directly against what we pre-theoretically supposed? To respond that, on this view, the actual facts concerning mental states are quite likely to be incompatible with our common sense conception of them is not at all reassuring. For it is difficult to see how this claim could be adequately justified.8

That completes my consideration of fully reductive brands of philosophical naturalism with respect to the mental. The next naturalistic project I wish to examine is of a more restricted character. Aspects of this project are to be found in the current work of Jerry Fodor.

III

Fodor has changed his views about what is needed to naturalize the mental from those expressed in the passage I quoted earlier in §I. For example, introducing his latest approach to intentionality, Fodor says:

... this will have the form of a physicalist, atomistic, and putatively sufficient condition... (1990b, p. 52) ... it's an attempt to solve Brentano's problem by showing that there are naturalistically specifiable, and atomistic, sufficient conditions for a physical state to have an intentional content... solving Brentano's problem requires giving sufficient conditions for intentionality, not necessary and sufficient conditions. (1990c, p. 96)

I shall examine carefully the connection between Brentano's problem and the issue of naturalization later in §IV. For the moment, it suffices to note that the problem, as Fodor sees it, is one of explaining intentionality in naturalistic terms. Such an explanation, we are now told, requires the production of appropriate sufficient conditions.

But what sort of naturalistic conditions is Fodor looking for? It is clear from his lengthy discussion of putative counter-examples that he is after conceptually sufficient conditions, ones that may be discovered by a priori reflection. This switch from providing a priori necessary and sufficient conditions for intentionality to providing just a priori sufficient conditions can be extended to cover the

8 For further compelling objections to the present version of naturalism, see Stich (1992).
non-intentional aspects of mental states too (although Fodor himself seems to wish to resist such an extension). What is now required to have a naturalistic theory with respect to pain or desire or belief is to have a priori naturalistic sufficient conditions for these states, that is, conceptually sufficient conditions couched in non-psychological vocabulary.

One can be a naturalist, then, on the above view, without having (or indeed supposing that there is) an illuminating reductive answer to Block’s question “What is common to all (actual and possible) pains in virtue of which they are pains?”. For the naturalistic sufficient condition met by actual creatures that are in pain in virtue of which they are in pain need not be met by pained creatures in other possible worlds. Indeed perhaps in some worlds there are no naturalistic conditions for pain. What matters to naturalism, on the present proposal, is how things are in our world, and whether the sufficient condition formulated by philosophical reflection is met by actual creatures. Let us call this more restricted version of naturalism “Conceptually Based Naturalism”. Now, concerning the search for sufficient conditions for a state to have a given intentional content, Fodor says, by way of further explanation, the following:

It’s enough if I can make good the claim that “X” would mean such and such if so and so were to be the case. It’s not also incumbent upon me to argue that since “X” does mean such and such, so and so is the case. (1990c, p. 96)

The project, then, for Fodor is to specify a condition C such that if C were to obtain, a state S would have the given content. C, moreover, must be formulated a priori in a vocabulary that eschews psychological and representational terms. Generalizing the project of Conceptually Based Naturalism is finding for each mental state type M (intentional or not), a naturalistic condition C which that if C were to obtain, M would be tokened.

One objection which has been raised to this project is that sufficient conditions can be produced at the drop of a hat once the requirement is removed that they also be necessary (see Stich 1992). Consider, for example, the state of believing that gorillas are placid. A sufficient condition for being in this state is that one be in a state that is expressible in English in the words “believing that gorillas are placid”. This condition (which uses no psychological vocabulary, although it harmlessly mentions some) is sufficient, since it is certainly true that were one to be in a state expressible by “believing that gorillas are placid” one would be in the given belief state.

Examples of this sort are surely unacceptable. For the doctrine of Conceptually Based Naturalism is intended to be a substantive philosophical position which, like Fodor’s latest proposal for naturalizing intentionality, is certainly not obviously true.

9 For the non-intentional aspects of mental states, Fodor seems to hold either that they have scientific functional essences or, more weakly, that sufficient conditions can be formulated for them in scientific functional vocabulary. Comments he makes in (1990c, p. 130) suggest that he presently inclines to the latter view.

10 This objection also occurred independently to me.
This objection is not compelling. Since our grounds for connecting belief states with their public linguistic expressions are clearly empirical, the sufficient condition stated above was not constructed by a priori reflection alone. Still, it is, I think, significant that what rescues the position of Conceptually Based Naturalism from the above objection is the requirement that the desired sufficient conditions be formulated a priori. For it is certainly not at all obvious that we have it within our power to construct a priori reductive sufficient conditions for either intentional or non-intentional mental states in non-mental, non-intentional vocabulary. Indeed, the past failure of philosophical attempts to produce analytic, reductive necessary and sufficient conditions in any common sense domain should surely make us wonder whether we are likely to be successful in conceptually elaborating sufficient conditions.

What follows if we fail in the case of intentional states, for example, according to the Conceptually Based Naturalist? Presumably, that intentionality cannot be naturalized (or at least that it may well be non-natural), and hence that Brentano’s Problem lacks an acceptable solution (or that it may well lack one). It is this thought, I take it, which motivates the view and makes the development of a satisfactory sufficient condition seem of the greatest importance (once the search for necessary and sufficient conditions has been rejected). The thought is mistaken, however. Intentionality, and mentality generally, do not need the truth of Conceptually Based Naturalism in order to be natural phenomena. So, the central motivation for the position is, I believe, suspect. In my view, the philosophical energy used in constructing putative sufficient conditions and testing them with thought experiments is misplaced, if one’s aim is to preserve a place for the mental in the natural realm. This will, I hope, become clear in the next section.

IV

Intuitively, naturalism with respect to the mental is, at root, the view that mental states are part of the natural world, just as much as chemical, biological, and geological states, for example. What is required to show that this view is true? Let us approach the issue indirectly by discussing some alternative ways of understanding the terms “mental” and “natural world”, beginning with the former.

According to one well known proposal, incorrigibility is the distinguishing feature of the mental (see Rorty 1970). That is to say, a report or description is classified as mental (or as being about a mental state) just in case it could not possibly be in error. This proposal derives from consideration of simple reports of conscious sensations (for example, “I am in pain”) in everyday contexts. However, it is clear that people do sometimes make mistakes about their mental states, as is shown by any number of psychological experiments (Nisbett and Wilson 1977). Furthermore, theories of cognitive psychology posit mental states that are inaccessible to consciousness. The claims made in such theories are empirical, and are therefore subject to revision and falsification.
Another proposal is that mental states or events are those to which their subjects have privileged access. This proposal admits the possibility of error for mental reports but insists nevertheless that the subject of the report is necessarily always in the best position to discover whether it is mistaken or to confirm it. So, for example, it is now conceded that I may mistakenly think at some given moment that I am having a visual experience of blue, say, but I have final epistemological authority on the matter: no-one else could be as well placed as I am to decide whether I am really having a visual experience of blue.

This suggestion is a minor improvement on the appeal to incorrigibility, but it is still indefensible. One problem, for example, is that subjects cannot have any privileged access to mental states which are inaccessible to consciousness.

A third proposal for characterizing the mental, due to Brentano (see Kraus 1919), is that mental entities are intentional. One important dimension of intentionality is the capacity to represent states of affairs or objects, whether or not the states of affairs actually obtain or the objects actually exist. Others, not stressed by Brentano himself but considered central by many philosophers today, are the capacity to represent only one of two co-instantiated properties, and the capacity to refer to or be about particular real things. Imagining something counts as mental by these criteria: one can imagine a three headed monster even though there are no such beasts. One can also imagine a creature to possess the property, having a heart, without possessing the property, having kidneys. Moreover, one can imagine with respect to a particular man, George Bush, say, that he is a Democrat. Likewise, wanting, hoping, doubting, and believing all pass the test. Unfortunately, some mental phenomena seem to fail. Consider, for example, having an itch or a tickle, or experiencing a pain. These mental states are not obviously representational at all.  

It appears no easy matter, then, to discover the essence of the mental. This should not overly concern us, however. For it seems to me unlikely that the mental has any interesting essence waiting to be discovered by philosophical reflection. I say this in part because none of the definitions proposed to date work, and in part because, as should be clear from the discussion in §II, the assumption that there is an illuminating list of necessary and sufficient conditions which characterizes the mental does not fit well with what we know of concept application elsewhere.

What, then, can we say about the mental? Well, whatever else the mental may be, it is a proper object of the study of the science of psychology, both cognitive and physiological. Cognitive psychologists are concerned, of course, only with our cognitive capacities. The laws that they formulate from experimental data typically make reference to the contents of the inner representational states that are hypothesized to constitute such capacities together with any associated behaviour. These laws are evidently ceteris paribus, since there are occasional conditions.

11 It is also not clear that all subdoxastic, representational, psychological states meet all three of the criteria for intentionality given in the text.
cognitive exceptions. Furthermore, the proper functioning of cognitive systems can be affected by breakdowns at the physiological level.

In the case of physiological psychology, the primary project is to uncover the underlying physiological bases (both within the brain and without) of the various types of cognitive and noncognitive mental states. Again laws are formulated from the experimental data. This time, however, the laws specify links between physiological states and conditions, mental states, and behaviour. For example, it has been proposed by Ungerleider and Mishkin (1982) that loss of the ability to recognize visually what kind of object is present without loss of the ability to identify where it is located is due to parieto-occipital damage. The reverse defect, it is hypothesized, is due to temporo-occipital damage. This defect manifests itself behaviourally in the failure of subjects to reach accurately for objects that have been correctly identified or to state reliably which one is on the left or further away.

None of the generalizations just stated are absolutely exceptionless, of course. Rather, like their counterparts in cognitive psychology, they have the status of ceteris paribus laws. This, it must be stressed, does not make them in any way unscientific. Ceteris paribus laws are the norm in science, not the exception. Consider, for example, the law of neurophysiology that if action potential is achieved, the neuron fires. This law admits of possible exceptions. Since the neuron typically fires at the end of the axon and action potential occurs at its base (the axon hillock), something can interfere with the electrical charge as it travels along the axon, thereby preventing the neuron from firing. Likewise, the law of genetics that when the genetic make-up of the human embryo is XY, it develops male sexual traits, is ceteris paribus. One exception here is provided by embryos with androgen insensitivity syndrome. Such embryos, even given the XY gene, develop female genitalia and other sexual characteristics. The same phenomenon is found in biological laws. For example, if a badger is attacked then, ceteris paribus, it fights back. So, science generally is riddled with ceteris paribus laws.

We are now ready to take up the question of how the term “natural world” is to be understood, and whether mental states are natural phenomena. According to some philosophers, to say that something is a part of the natural world, as we normally conceive it, is to say that it participates in causal interactions which fall under scientific laws and theories. Glaciers, for example, are part of the natural world, on this view, since they enter into causal interactions described in the science of geology. Similarly, fossils: in this case, the science is paleontology.

Now, on this understanding of “natural”, it is evident that mental states are natural phenomena. For they are the subject matter of scientific laws, namely those comprising the science of psychology. Unfortunately, naturalism, so understood, is a doctrine not worthy of the name. The immediate problem is that it does not rule out the possibility that mental states are states of Cartesian souls that always inhabit some living body or other. Such states enter into causal interactions that are describable by psychological laws. Yet they are surely inimical to the naturalist position.
One way of handling this problem is to maintain that the natural world is better thought of as consisting, at the level of particulars, only of spatio-temporally localized items that meet the condition specified above. Properties and states count as natural phenomena, on this account, so long as their tokens are localized in the spatio-temporal net and they themselves are projected by scientific laws. This evidently eliminates Cartesian souls from the natural world. But it does not extirpate all difficulties, I suggest. In particular, it still allows the natural world to contain spatio-temporally localized spirits or ghosts with no intrinsic properties other than their psychological ones. And this is surely unacceptable on any straightforward understanding of naturalism.

At this stage, it might be argued that since Cartesian souls and spatio-temporally localized spirits of the above sort are certainly not physical things, we should simply take the natural world to be the physical world, so that something (token or type) is counted as natural if, and only if, it is physical. Many philosophers would reject this proposal on the grounds that it makes naturalism about the mental almost certainly false. That is not my own view, however. It seems to me that before we can evaluate the consequences of the above proposal for psychological naturalism, we need to consider what the crucial term “physical” means. It is sometimes supposed that a general term is physical (that is, that it picks out a physical state or property or kind) just in case it occurs in some true theory of physics. This is evidently too narrow a definition: terms like “acid”, “alkali”, and “DNA” lie outside the domain of physics yet would normally be classified as physical. Perhaps we should say, then, that a general term is physical just in case it occurs in some true theory of physics, chemistry, molecular biology, or neurophysiology. But it is far from clear that this is a satisfactory way to characterize the physical. If “gene” and “neuron” are now classified as physical terms, then why not go further and classify “tse-tse fly”, “crocodile”, “continent”, and “planet” (terms found in entomology, zoology, geology, and astronomy respectively) as physical too? The general problem here, of course, is that we have not been provided with any account of what physics, chemistry, molecular biology, and neurophysiology share in virtue of which they count as physical and the other sciences mentioned above do not.12

One way of avoiding this problem is to say that a general term is physical just in case it occurs in some true theory adequate for the explanation of the phenomena of non-living matter. But there remain serious difficulties even here. Suppose that there are properties that are tokened only in the brains of certain living creatures, and that these properties figure in neurophysiological laws. It seems to me ad hoc to deny that such properties are physical (Block 1980b, p. 296). Yet this is what we must do according to the final definition.

How, then, is the term “physical” to be understood? If this question is taken to demand a fixed list of necessary and sufficient conditions for the application of “physical”, as it is ordinarily used, then it should be no surprise to hear that I very much doubt that it has an answer. As I have repeatedly emphasized, neces-

12 This problem is discussed by Carl Hempel in his (1970).
sary and sufficient conditions are hard to come by for any terms, let alone ones at this level of abstraction. Perhaps the best we can say is something like this: Physics is the paradigm or prototype for the physical sciences. A given science counts as physical, then, so long as it is sufficiently similar to physics, and a given general term counts as physical so long as it occurs in a physical science. The notion of sufficient similarity at work here is vague and multi-dimensional. Thus, conflicts may arise about whether to count “glacier”, say, as a physical term because different competent users of “physical” may rely tacitly on different dimensions of similarity.

So where does this leave us with the question of the status of mental state types? As I noted earlier, whatever else the mental may be, it is a proper object of study of the science of psychology (both cognitive and physiological). So, assuming that a general term counts as physical just in case it occurs in a physical science, whether we count mental states as physical depends upon whether we classify psychology as a physical science. Given my comments above, this seems to me to be an issue on which there is no single, fixed, correct view. But it seems to me that geology is reasonably classified as a physical science. And if geology is, why not psychology?\(^{13}\) The conclusion I draw, then, is that mental state types may reasonably be taken to be physical. So, naturalism with respect to the mental is not directly threatened by the latest understanding of “natural world”.

Regrettably the problem with which we began has not really been solved. This is because the mental states of Cartesian souls and spatio-temporally located spirits with only mental properties are not clearly non-physical on the present proposal. After all, ex hypothesi, they are referred to in psychological laws, and psychology, we are now supposing, may reasonably be classified as a physical science.

One might respond that if psychological laws do refer to such items then psychology is, on that account alone, sufficiently dissimilar to physics not to count any longer as a physical science. This seems to me reasonable enough; but in the present context it has the effect of throwing the baby out with the bath water. For if psychology is not classified as a physical science then, given the above proposal, naturalism with respect to the mental is not true.

Just how is naturalism with respect to the mental to be explicated? The answer, I suggest, is that, for the naturalist, mental states should not only participate in causal interactions describable in scientific laws but also bear the same general ontic relationship to lower level physical items as do the physical entities quantified over and referred to in higher level physical laws generally (for example, those in biology and geology). I want now to say something about what I take this relationship to be.

\(^{13}\) The need to consider other special sciences in assessing the status of psychology has been emphasized by Fodor more than any other philosopher. See e.g., his (1990a). However, if what I say in the remainder of this section is correct, Fodor did not take his own advice sufficiently to heart.
Consider Mount Everest. Suppose that \( t \) is one of the chunks of matter Everest would have lost had certain bombs been detonated at its top. Suppose also that the bombs are not in fact detonated. Then it is true that Everest might have existed without \( t \). But the same is not true of the mereological sum of Everest’s parts. Hence, Everest and the sum differ in a modal property. Hence, Everest is not strictly identical with this sum. Rather, Everest is constituted by it. What is true here for Everest is true for other geological objects. A glacier, for example, is constituted by a massive chunk of ice, which, in turn, is constituted by an aggregate of water molecules. Indeed, the same is true for higher level physical objects and events generally.\(^{14}\) Each horse is constituted by a torso, four legs, a head, a tail. Each horse leg is constituted by a thigh, a calf, a hoof. Likewise, a predator’s eating its prey is constituted by an action of chewing and swallowing. Each action of chewing is constituted by certain movements of the jaw, and each action of swallowing by certain movements in the throat. In all of these cases, modal considerations show that there is no strict identity.

For the naturalist, psychological objects, processes, and events should be no different. Consider, for example, mental images. According to one prominent cognitive scientist, Stephen Kosslyn, each mental image is a pattern of filled cells in a special medium called “the visual buffer” (Kosslyn 1980, Tye 1991). The “is” here, I maintain, is the “is” of constitution, not identity. The argument for this claim is straightforward. Images are subject to what Kosslyn calls “transformation processes”—processes such as rotation and scanning. What these processes do, on Kosslyn’s theory, is to alter the pattern of filled cells in the visual buffer. They do not, however, destroy the mental images on which they operate. When one scans across an image, for example, one still has the same image as before. Likewise, when one rotates an image in two dimensions, one surely does not have, after rotation, a second image. After all, one would certainly not describe oneself here as having had two images. The rotated image is the same image as before: it merely represents the given object at a different orientation.

Consider, then, a particular mental image \( M \). Suppose that \( M \) is not in fact subject to any of the above transformation processes. Still, it might have been. Had this occurred, some of \( M \)’s component cells would have changed their contents and the overall pattern of filled cells would have been a different pattern. Hence, \( M \) has a modal property that the pattern lacks, namely, the property of possibly having such and such cells differently filled. Hence, \( M \) is not strictly identical with the pattern. Rather, \( M \) is constituted by it. Hence, given that each pattern is physically located in the brain, the relationship the relevant brain processes bear to \( M \) must also be one of constitution.

It seems to me plausible, then, to suppose that the naturalist perspective requires only that mental tokens generally be constituted by neural processes, just as neural processes, in turn, are constituted by molecular processes. If mental tokens are viewed in this way, then it must be granted that each such token may vary in its constitution in different possible worlds. Something similar to this is

\(^{14}\) For further discussion here, see Boyd (1980).
true of mental types, I believe, if the naturalist perspective is adopted. Let me explain.

In general, higher level physical types are not identical with lower level ones. There are no types from chemistry and physics, for example, with which being a neuron, being a continent, being an earthquake may plausibly be identified. Moreover, as Paul Churchland (1984, p. 41) has noted, not even the property of temperature—a property often cited by reductionists as the model for reduced properties generally—is really identical with any lower level property. True, temperature is mean molecular kinetic energy, but only in a gas where the molecules are free to move around. In a solid, temperature is something different, since the molecules are confined to certain restricted periodic motions. In a plasma, there are no constituent molecules, but there is temperature. Finally, in a vacuum, there is "blackbody" temperature in the distribution of electromagnetic waves, although there is certainly no molecular kinetic energy.

The general relationship which obtains between higher level and lower level physical properties is one of realization: temperature is realized by mean molecular kinetic energy in a gas and by the blackbody distribution of electromagnetic waves in a vacuum. The realization relation is, at least in part, one of determination: the lower level property synchronically fixes the higher level one, so that the tokening of the former at any time $t$ necessitates the token of the latter at $t$ but not conversely. (I take the relevant notion of necessity here to be both nomological and strict—that is, without any ceteris paribus qualification.)

The parallel between types and tokens on the above conception of naturalism should now be clear: higher level types may be realized by more than one lower level type within the actual world; higher level tokens may be constituted by different lower level tokens but only in different possible worlds.

We are now in a position to summarize what naturalism with respect to mental states (token and type) comes to on the above account:

Mental states participate in causal interactions which fall under scientific laws, and are either ultimately constituted by or ultimately realized by microphysical phenomena.

The term "ultimately" appears here, since, in the naturalist perspective, there is a hierarchy of constitution and realization relationships between higher level and lower level physical items which has as its foundation the microphysical realm. So, the constitution and realization of the mental by the microphysical is certainly not itself an unexplained brute fact. Rather, it is explained via the intermediate stages of the hierarchy. Non-intentional mental state types, for example, are realized by types from microphysics in virtue of the latter realizing chemical types, which, in turn, realize neurophysiological types, which finally realize the relevant mental types.

The view just summarized is, I believe, consonant with our pre-theoretical understanding of naturalism with respect to the mental. For surely, as I remarked at the beginning of this section, the key idea in naturalism is that the mental is a
part of nature in the same way as the chemical, biological, and geological. And that idea is not only preserved but also sharpened.

Why believe such a view? The answer is simple: the mental is studied by psychology. Psychology is a science no different in its procedures and laws from other sciences. So, of course, the mental is a part of nature in the ways I have described. To suppose otherwise is to suppose that there is something peculiar about the mental which prevents it from having the features adumbrated above. And there just is no good reason for any such supposition. So, as I commented in the introduction to this paper, naturalism with respect to the mental, once properly explicated, is really beyond question.

I have focussed so far in this section on the status of mental states without saying anything specifically about their intentional aspects. Some philosophers opposed to naturalism may see here a straw at which to grasp. However, the points I have made apply just as well to intentional mental states as to non-intentional ones. Since the causal interactions of intentional states lie in the domain of cognitive psychology, there are psychological laws that advert to intentional content (for example, the law that believing that \( P \) and \( Q \) brings about believing that \( P \), ceteris paribus, or the law that if people are asked questions about properties or aspects of things they have seen, and they have not thought about those properties or aspects much, then, ceteris paribus, they form images of the things before answering). So, the intentional aspects of mental states figure in causal transactions that fall under scientific laws. There is also no more reason to deny that they are realized by lower level physical states than there is to deny that the non-intentional aspects of mental states are so realized. Of course, in this case, given the world-involving character of intentional content, the relevant physical states will include external ones. But so what? Biological functions, for example, the function possessed by the bee’s dance of assisting other bees to find nectar, are evidently not determined by purely internal physical properties of the bearers of the functions. After all, the functions are typically specified via relations to environmental items; and two items made of the same materials may have different functions in different environments (as, for example, in the case of the tails of possums and lions). Still, this obviously does not show that biological functions lack any lower level physical realizations. Rather, all it shows is that the physical realizations themselves must involve external items. The same is true for intentional mental states.

I want now in closing to make some remarks on naturalism and Brentano’s problem. We saw earlier that Brentano was struck by the fact that we can, it

---

15 Intentional content can be world involving in two ways: external items, abstract and concrete, can enter into it; or, more narrowly, items, abstract and concrete, in the immediate environment can enter into it. The latter claim, as it is normally understood, is stronger than the former, since it requires that the relevant external items be appropriately causally connected to intentional state tokens. Colin McGinn has labelled the two views “Weak Externalism” and “Strong Externalism”. I assume that one or other of these two views are correct, and hence that purely internal conceptions of content are mistaken, but I shall not attempt to adjudicate between them in the present paper. For more on the two varieties of externalism, see McGinn (1989).
appears, think about things which do not exist, for example, golden mountains. How are we to explain this fact? Brentano’s view was that it reflects a feature of thought, namely intentionality, which cannot be reduced to anything in the natural realm. He then went on to argue from this that mental states generally cannot be part of the natural order.

Contemporary naturalists of a reductionist bent take Brentano’s problem to require them to frame conditions for intentionality in naturalistic terms. This, it seems to me, is a mistake. On the view I have presented, intentionality is already naturalistic. If neither sufficient nor necessary and sufficient conditions can be formulated by philosophers in non-intentional vocabulary for a state to have an intentional content, what this shows, contra Brentano, is that intentionality cannot be analyzed in terms of anything else in the natural realm.

What I am suggesting, then, is that both Brentano and naturalists such as Fodor have wrongly supposed that unless we can specify appropriate a priori conditions for thinking about X’s which are naturalistic and which clearly, unproblematically do not require that there really be any X’s, we have reason to believe that the intentionality of thought is a non-natural phenomenon. But what we should believe in this event is that the capacity to think about things which do not exist is a natural phenomenon whose operation we cannot further explain by conceptual reflection.

Once it is granted that intentional vocabulary is itself naturalistic, however, the way is opened to a solution to Brentano’s problem which incorporates intentional terms. In particular, it now makes sense to try to explain the capacity to think about golden mountains in terms of the capacity to have thoughts, the intentional contents of which contain the property of being a golden mountain. For the latter thoughts do not require that there actually be any golden mountains. Whether such an explanation is ultimately defensible is an issue I cannot pursue here. But to attack it on the grounds that it is not sufficiently reductive is to miss the point.

It should be clear by now that the conception of naturalism I have been elucidating requires none of the brands of reductive naturalism I rejected earlier. Indeed, on my account, there is really no more motive for trying to develop conceptual analyses or conceptually sufficient conditions or type identities in defence of naturalism with respect to the mental than there is in defence of naturalism with respect to the seismological or the entomological. Type reductions are not needed in the former case any more than they are in the latter. This is, I

Likewise for the other features of intentionality adumbrated earlier.

Instances of this property are not found in the environment. However, the property itself is an external entity, and its component properties (being golden, being a mountain) are environmentally connected to thinkers. So, neither Weak Externalism nor Strong Externalism, on any version which requires only that the properties entering into intentional contents be decomposable into environmentally located properties, is challenged by any claim in the text.

Nor can I take up questions concerning the other aspects of intentional content I distinguished earlier except to comment that it seems to me that illuminating, similarly non-reductive, explanations can be given of these aspects on the supposition that intentional states have a linguistic structure.
believe, all to the good. For the prospects for successful reductions of the above sort are bleak. Naturalism concerning the mind is, or should be, a very natural, straightforward view. What has been unnatural, I suggest, is much recent philosophical theorizing on its behalf.\(^{19}\)

---

REFERENCES


\(^{19}\) In its negative aspects (i.e., in its opposition in §§1 - III to certain type level naturalistic projects), this paper is very similar in spirit to Stich (1992). The views expressed in these sections were, however, developed independently, and the paper submitted to *Mind* only a few weeks after his. The present version of the paper has been revised to take account of Stich’s article, and, in order to avoid excessive overlap, some material has been excluded (e.g. a lengthy discussion of prototypes). Stich tells me it is a novel experience to find anyone who actually agrees with him—a sure sign of middle age. For my own part, I am equally surprised. Whether our convergence on these matters is an indication of Stich’s moving toward the middle or my suddenly turning radical, I leave for the reader to decide.


Sober, E. 1985: “Putting the Function back into Functionalism”. Synthese, 64, pp. 165-93.


