

Functionalism

David Papineau

Philosophical Concept

The term 'functionalism' means different things in many different disciplines from architectural theory to zoology. In contemporary philosophy of mind, however, it is uniformly understood to stand for the view that mental states should be explained in terms of causal roles. So, to take a simple example, a functionalist in the philosophy of mind would argue that pains are states which are normally caused by bodily damage, and tend in turn to cause avoidance behaviour. Functionalism is often introduced by an analogy between mental states and mechanical devices. Consider the notion of a carburettor, say. For something to be a carburettor it need not have any particular physical make-up. Carburettors can come in many different materials and shapes. What makes it a carburettor is simply that it plays the right causal role, namely that it mixes air with petrol in response to movements of the accelerator and choke. Similarly, argue functionalists, with the mind. The possession of mental states does not depend on the physical make-up of the brain; it depends only on its displaying the right causal structure. Since organisms with very different sorts of biological make-up, like octopuses and humans, can have states with the causal role of pain, say, it follows from functionalism that octopuses and humans can both be in pain.

There exists a number of different subspecies of functionalism. One important division depends on how the relevant causal roles are determined. 'Common-sense' functionalists take them to be fixed by common-sense psychology; 'scientific' functionalists take them to be fixed by the discoveries of scientific psychology. So, for example, common-sense functionalists will hold that emotions play the causal role that common-sense psychology ascribes to emotions, while scientific psychologists will argue that scientific psychology identifies this causal role. Functionalism, of whatever subspecies, is open to a number of well-known criticisms. One central objection is that it cannot accommodate the conscious, qualitative aspect of mental life. Could not a machine share the causal structure of someone who was in pain, and thereby satisfy the functionalist qualification for pain, and yet have no conscious feelings?

It might seem that functionalists can respond to this difficulty by being more stringent about the requirements involved in the causal role of a given human sensation. But there is a danger that functionalism will then lose much of its appeal. The original attraction of functionalism was that its 'liberal' specification of causal roles allowed that humans could share mental states with non-

humans. This feature is likely to be lost if we switch to more 'chauvinist' specifications designed to explain why non-humans do not share our conscious life.

Another objection to functionalism is that it cannot account for mental representation. Functionalism focuses on the way mental states enter into causal structure. But it is doubtful that mental representation can be explained in purely causal terms.

Some philosophers argue that the issue of mental representation can be dealt with by adding some teleology to functionalism, that is by considering the biological purposes for which mental states have been designed, as well as their actual structure of causes and effects. However, once we do appeal to teleology in this way, it is not clear that we still need a functionalist account of representational states, for we can now simply identify such states in terms of their biological purposes, rather than their causal roles.

1. Origins of functionalism

Originally functionalism was a response to philosophical behaviourism. The behaviourists rejected the traditional Cartesian picture of the mind as an essentially private realm accessible only to the conscious subject (Descartes 1641). Instead they argued that mental states are dispositions to behaviour, and so are publicly accessible. To desire an ice cream, said the behaviourists, is to be disposed to eat one when the opportunity presents itself (see Behaviourism, analytic; Private states and language).

The functionalists argued that behaviourism fails to distinguish sufficiently between mental cause and behavioural effect. They argued that there is no simple pairing of mental states with pieces of behaviour, since which behaviour issues from any given mental state will depend on the agent's other mental states. My desire for an ice cream will make me walk to the fridge if I believe it contains an ice cream, but it will make me walk down the street to the shop if I believe I can buy an ice cream there.

Because of this, functionalists argued that mental states are inner causes distinct from their behavioural effects. In saying this, however, they did not want to return to the Cartesian conception of mental events as essentially private states. In their view, mental states are part of the public world of causes and effects studied by science. They are 'inner' only in the sense that they are unobservable causes of overt behaviour, in the same way as atomic structures are unobservable causes of chemical reactions.

Because they take mental states to be unobservable, functionalists think that we only have an indirect grasp of their nature, as playing a certain causal, or 'functional', role in a cognitive system. One important consequence of this is that different physical states might play the relevant role in different beings, or even in the same being at different times. (Indeed

functionalism as such leaves it open that this role could be played by some special non-physical state. In this sense functionalism is compatible with dualism (see Dualism). However, most functionalists also hold, for independent reasons, that mental roles are in fact filled by physical states. I shall simplify the following discussion by adopting this assumption.)

Some early functionalists, most notably Hilary Putnam (see Block 1980), appealed to an analogy with computers to add precision to the idea of a causal role. Putnam pointed out that any programmed computer can be abstractly characterized as a Turing machine, independently of its 'hardware' or physical make-up; he then argued that any two systems will share mental states as long as they have the same Turing machine description (see Putnam, H. §6; Turing machines). Today, however, it is more common to elaborate functionalism in terms of the analogy with scientific unobservables.

Unobservable entities of any kind pose a *prima facie* problem which has been much discussed in the philosophy of science. How should we understand terms like 'ionized', 'radioactive', 'diatomic', and so on, given that we have no direct access to their referents? Frank Ramsey argued that our grasp of such terms derives from our theories about the relevant unobservables. If we have a theory $T(P_1 \dots P_n, O)$, about how various unobservable properties, $P_1 \dots P_n$, relate to each other and to observables O , then we can read claims involving these properties as claims about whichever properties happen to play the relevant theoretical roles. More precisely, a claim attributing property Π , to individual a , say, can be read as the claim $(F_1 \dots F_n)(T(F_1 \dots F_n, O) \& F_i a)$. (In words: 'There exist properties $F_1 \dots F_n$, which are related to each other and to observables as T says, and a has the i th one.') Note that this way of understanding claims about the P s does not credit us with any prior understanding of these terms, but only with an understanding of existential quantification and of the observable terms O (see Ramsey, F.P. §5).

Functionalists take an analogous attitude to terms for mental states, like 'belief', 'desire', 'jealousy', 'pain', and so on. Suppose that our psychological theory contains such assumptions as that: anybody faced with an ice cream will believe there is an ice cream in front of them; anybody who is hungry and hot will desire an ice cream; anybody who desires an ice cream and believes one is in front of them will reach out for it; and so on. Then we can understand ascriptions of desires and beliefs to particular people as claims that there are states which behave as this theory claims, and that the people in question have them.

2. Common-sense and scientific functionalism

There are various brands of functionalism. One division depends on whether the theory T used to introduce mental terminology derives from common-sense psychology or from scientific research. This issue also determines whether functionalism can claim to give an account of the *meaning* of everyday mental terms like 'desire' and 'pain'.

If the theory T is derived from common-sense psychology, then it is open to functionalists to argue that the resulting Ramsey-style account of mental discourse explicates what everyday people mean by terms like 'desire' and 'pain'. That is, they can argue that these terms, in everyday discourse, simply signify 'the causal roles specified by T'.

On the other hand, if the theory T used to introduce mental terminology is not part of everyday thinking, but rather some new scientific theory, then it cannot plausibly be argued that everyday thinkers derive their grasp of terms like 'desire' and 'pain' from this theory. So functionalists who want to replace common-sense psychology by some new scientific theory need to view the Ramsey-style account as fixing the meaning of new technical terms, which signify the causal roles specified by the new scientific theory. It is then open to them to argue that the states picked out by their technical terms are in fact the same states as referred to by the everyday terms 'desire' and 'pain'. But, from the point of view of this kind of scientific functionalism, this will be a synthetic matter, to be supported by empirical evidence, and not a matter of definition.

The former kind of functionalism, which I am calling common-sense functionalism, is sometimes also called 'analytic' functionalism, in recognition of the fact that it makes it a matter of definition that everyday mental terms stand for certain theoretical roles. It is perhaps worth pointing out, however, that the common-sense theory T involved in these definitions will not itself be analytic. What is analytic is this claim: *if* there are states which play the roles specified by T, *then* these states are desires, pains, and so on. But the further claim, that there are in fact such states linking sensory inputs with behavioural outputs, is obviously synthetic.

One issue facing both common-sense functionalists and scientific functionalists is how *much* of the relevant theory to count as contributing to the meanings of mental terminology. It is obviously unsatisfactory to include all the many assumptions of everyday or scientific psychology in our Ramsey-style definitions of mental terms. For it would then follow that, if any single one of these assumptions turns out to be false, then all these terms will fail to refer to anything (since there will not then be any states which play the precise causal roles specified by the original incorrect theory).

The obvious remedy is to argue that only some core set of assumptions from the relevant theory plays a part in fixing the meanings of its terms. This would in effect make it an analytic requirement that desires, or pains, or whatever other mental states are at issue, satisfy the relevant core assumptions, while leaving it synthetic that they satisfy any further assumptions. However, there are general doubts about any sharp analytic-synthetic distinction of this kind. What principled reason could there be for picking out certain theoretical assumptions as constitutive of the meaning of 'desire', say, while excluding others? Perhaps the best solution for the functionalist would be to postulate some more-or-less vague distinction between the central and non-central assumptions of the relevant theory, and then take it to be a matter of meaning

that the various mental terms will satisfy most of these assumptions. This will admittedly make it a vague matter whether the relevant mental terms apply to states which do not satisfy all the central assumptions; but this vagueness is arguably a feature of all theoretical terminology in science, and not necessarily vicious.

3. Roles and realizers

A further division among functionalists depends on whether they identify mental states with 'roles' or 'realizers'. Should we equate pain, say, with the property of having-some-state-which-plays-the-requisite-causal-role; or should we equate it with whichever physical state in fact realizes that role? Suppose, to adopt the usual philosophical oversimplification, that the pain role is realized in human beings by C-fibres firing. The question is then whether our term 'pain' picks out the 'realizer' property *C-fibres firing*, or whether it picks out the 'role' property *having-some-physical-state-which-plays-the-pain-role*. Note that the role answer implies that humans and octopuses, say, share the same mental state when they are in pain, for it is true that they are both in a physical-state-which-plays-the-pain-role, even if octopus pains are realized by a quite different physical state from C-fibres firing. The realizer answer, by contrast, implies that in this case human pains are different states from octopus pains, since the physical states which realize these roles are different.

The issue here is whether mental properties are first-order properties or second-order properties. By way of comparison, consider the property of being coloured. This is a second-order property, since you have it if you have some first-order property (red, blue,...), which in turn has the property of being a colour. Role functionalists maintain that mental states are second-order properties in this sense, in that you have them if you have some first-order physical property, which in turn has the property of playing a given causal role.

The obvious argument in favour of the role answer is precisely that it does allow humans and other beings to share pains. It seems odd to hold that the state of pain can only be present in those beings that possess C-fibres. Indeed one of the original attractions of functionalism was that, by detaching mentality from physical make-up, it seemed to allow beings of quite different physiologies to share the same mental state.

Those functionalist philosophers who favour the realizer view will allow that there is a role property common to humans and other beings with different physiologies, and that we understand the concept 'pain' by associating it with this role property. But they nevertheless argue that this term still refers to different realizer states in application to the two species. By analogy, consider the word 'eye'. All eyes have a common role feature (namely, that they are sense organs which respond to visible light), in virtue of which they are all eyes, despite their different physical realization in different species. Nevertheless, when we use the term 'eye' in

application to some individual organism, we seem clearly to use it to refer to a physical part of that individual, not to the individual's instantiation of some abstract property.

In the end the debate between role and realizer versions of functionalism hinges on whether we take physical instantiation or causal role to be the essential feature of states like pain. Does difference in physical instantiation, or only difference in role, imply that we have a different state? This may seem an overly nice metaphysical issue. But it matters to the question of whether functionalism supports the identity theory of mind. The identity theory argues that mental states are identical with physical states (see Mind, identity theory of). The realizer version of functionalism agrees with this identity theory, since it identifies human pains, say, with the physical state of C-fibres firing. But role functionalism denies the identity theory, since it equates human and other pains with the second-order role states.

Of course, this does not mean that role functionalism is not 'physicalist' in some broader sense. For the role properties it identifies with mental properties are still *realized* by physical properties. We should not think of the role properties as akin to substantial non-physical properties, like the properties a Cartesian dualist would ascribe to mental substance. Rather, they are simply second-order properties – in particular second-order properties that are guaranteed by the possession of physical properties with certain causal roles.

4. Inverted spectra

Now to criticisms of functionalism: one family of objections focuses on functionalism's ability to deal with conscious, qualitative states – states that it is 'like something to have', in Thomas Nagel's 1974 phrase. These states include sensory experiences, pains, itches, and emotions, and so on, but arguably do not include such propositional attitudes as belief and desire (see Qualia; Propositional attitudes.)

Let us start with the 'inverted spectrum' argument against functionalism. Suppose that baby Matthew has an operation performed on his retina at birth which switches the 'red' and 'green' messages from his retina to his visual cortex: the central physiological state produced in Matthew by red things is thus the state normally produced by green things in other people, and vice versa. From then on Matthew is brought up normally, learning how to discriminate between red and green things, to call them 'red' and 'green' respectively, and so on. Consider now the state (let us call it A) produced in Matthew by red things. It seems likely, given Matthew's normal upbringing, that A will play the same causal role as the physiologically different state which arises when other people are presented with red things. Now let us ask: what kind of conscious experience will Matthew have when he is in state A? Despite A's sharing a causal role with the state normally produced by red things, it seems intuitively plausible that A will be consciously like the state normally produced by green things. After all, when Matthew has A, his brain is physiologically just like a normal brain presented with a green thing.

But this now presents a *prima facie* problem for functionalism. For if the functional role of A classifies it with the state normally produced by red things, but it feels like the state which is normally produced by green things, then functionalism has failed to capture the conscious aspect of this state.

Note how this thought experiment differs from the simpler, traditional ‘inverted spectrum’, which postulates an individual – Millie, say – who is normal in all physical respects but still has her colour experiences ‘inverted’. The physical state which gives normal people the conscious experience of red gives her the conscious experience of green. Millie is incompatible with the general physicalist assumption that there cannot be mental differences without physical differences of some sort. The ‘retinal operation’ version of the inverted spectrum thought experiment, by contrast, does not require us to deny this general physicalist assumption. There are physical abnormalities in Matthew to account for his abnormal colour experiences. The point of the retinal operation thought experiment is rather that Matthew’s abnormality is only at the physical level, and not at the functional level. Because of this, the retinal operation thought experiment poses a problem specifically for functionalism, which is committed to explaining mental differences in terms of functional differences, but not necessarily for other versions of physicalism.

5. Responses to inverted spectra

There are various responses to the retinal operation thought experiment open to both ‘role’ and ‘realizer’ functionalists. Let us consider them in turn, starting with role functionalism.

(1) The most direct response open to role functionalism is simply to deny the intuition that Matthew will have different conscious experiences from normal people. That is, role functionalists can argue that if Matthew’s state A plays the same functional role as the state produced in normal people by red things, then it will feel the same, even if it is physiologically different. After all, they can point out, by hypothesis this state will make Matthew react in just the way that normal people react to red experiences, both in his behaviour and in forming further beliefs and desires. If this is so, they will ask, then what substance is there to the hypothesis that the experience is nevertheless consciously different?

However, this answer is less than wholly convincing. Maybe some qualitative states, like pain, can plausibly be argued to depend on nothing but functional role (if a state creates a pressing and intense desire to move some part of your body, does that not show it is a pain?). But other qualitative states, of which colour experiences are the paradigm, do not seem to be nearly so closely tied to functional role (there is no specific desire or behaviour which is typically prompted by an experience of green). So if we take some such state, like normal people’s experience of green, and imagine its functional role switched while its physiology stays the

same, as in Matthew, then most philosophers have a strong intuition that will still feel the same, despite the fact this runs counter to role functionalism.

(2) Other role functionalists adopt a different tack. They simply admit that their theory of mind cannot deal with 'qualia' – that is, with the qualitative aspects of colour and similar conscious experiences. Instead they accept that these aspects are fixed by physical realization rather than functional role. Role functionalists of this stripe can continue to maintain a role functionalist account of non-qualitative states like belief and desire, and indeed of the non-qualitative aspects of qualitative states. But their role functionalism will be incomplete, in that they admit that qualia themselves are not fixed by their functional role. In particular, they concede that an individual who is functionally normal but physically abnormal, like Matthew, will not experience normal qualia.

(3) The 'retinal operation' thought experiment is much less of a problem for realizer functionalism than for role functionalism. After all, realizer functionalists take mental terms to refer to physical states from the start: they distinguish between human pains (one physiological state) and octopus pains (a different physiological state), even though both states play the pain role. Similarly, they can distinguish between Matthew's experience of red, which is one physical state, and a normal person's experience, which is a different physical state. True, both states play the same functional role. But nevertheless the realizer functionalist counts them as distinct mental states, and to this extent need not be disturbed by intuitions that they are also qualitatively distinct.

(4) Both the realizer functionalist of the last paragraph, and the incomplete role functionalist of the previous one, take qualia to be fixed by physical realization. Different physics, different qualia. However, some philosophers argue that this principle is excessively 'chauvinist', since it implies that beings who lack human physiology cannot share human experiences. Surely, more 'liberal' philosophers argue, we want to allow that dolphins, say, or the inhabitants of Proxima Centauri's third planet, might be functionally organized in such a way as to experience pain and sadness, say, even if they have non-human physiologies.

David Lewis has devised an ingenious version of functionalism, which allows this kind of liberalism while still respecting the intuition that Matthew's experience when he is faced with something green is like a normal person's experience of red. Lewis argues that terms for mental states ('belief', 'desire', 'pain', 'sadness', 'experience of green' and so on) should be understood, in their application to any given being, to refer to the first-order state that realizes the relevant functional role *in normal members of the species* (or other group to which the being belongs). Matthew is a human being, so 'experience of green' as applied to him refers to the physical state that is produced by green things and plays the corresponding functional role in normal human beings. In Matthew, of course, this physical state (state A) plays the functional role that relates to

red things. But classified mentally it is still an experience of green, because in normal people it plays the role appropriate to green things.

If we are talking about a Proxima Centaurian, by contrast, then 'experience of green' does not refer to the physical state produced by green things in humans, but rather to the physical state (if there is one) produced by green things in Proxima Centaurians. So it is possible for a Proxima Centaurian to have an experience of green, Lewis argues, as long as it is in the physical state that plays the relevant role in its normal conspecifics, even if that is different from the state A that plays this role in humans. (Note that there could also be an abnormal Proxima Centaurian in whom the experience of green does not play the green functional role, as long as it is the state that plays that functional role in normal Proxima Centaurians.)

So, according to Lewis, qualitative experiences vary with physical realizer states within a given group, but with functional role states across groups. His theory thus accommodates the intuition that Matthew is qualitatively different from normal humans, while avoiding the chauvinist implication that animals and extraterrestrials cannot share our experiences.

Lewis is normally classified as a realizer functionalist. But in one respect the mixed theory just outlined goes beyond realizer functionalism. A different physical state realizes experiences of green in humans and Proxima Centaurians. Yet Lewis counts them alike in respect of their qualitative nature. This makes him different from the realizer functionalists described in (3) above, who take the difference in physical realization to lead to a difference in qualitative feel. Still, this only shows that Lewis is not a straightforward realizer functionalist, not that he is wrong.

A more serious difficulty for Lewis' theory is that it makes the qualitative classification of peoples' mental states depend on which species or group we assign them to. Suppose a minority subspecies of humans evolves a different pain mechanism from other humans. And suppose Jane belongs to this subspecies. *Qua* member of the subspecies, Jane will be in pain when her mechanism is activated, for she will be in the state that plays the pain role in normal members of the subspecies. But *qua* human *simpliciter*, she will not be in pain, for she is not in the state that plays the pain role in normal humans. Lewis accepts that there will sometimes be no unique answer to the question of which group an individual belongs to, and consequently that it will be indeterminate what qualitative mental states it has. This is, to say the least, a surprising consequence of his theory.

(5) Some role functionalists argue that the way to deal with the inverted spectrum problem is to distinguish extra levels of functional organization. Standard discussions only consider two levels – 'macroscopic' functional role and physical realization. Because of this, standard discussions face a dilemma: they either tie qualia to 'macroscopic' functional role, and fail to accommodate

the intuition that Matthew is qualitatively abnormal; or they tie qualia to physical realization, and end up 'chauvinistically' denying pains to Proxima Centaurians. Lewis' mixed theory offers one way out. But another way out would be to uphold role functionalism, but identify some intermediate level of 'micro-functional' organization, which will distinguish Matthew from other humans, yet be common to humans and members of different species. By hypothesis, Matthew's states coincide with normal human states in respect of such 'macroscopic' role features as which actions they give rise to, which beliefs and desires they prompt, and so on. But they need not coincide in more 'microscopic' respects such as the light reflectance profile being computed, the firing pattern of the relevant neural units, and so on. The important point is that these microscopic features are still functional, in that they can be realized in systems of different physical composition. So they could be shared by humans and physically different species, while at the same time serving to distinguish Matthew from normal humans.

Note that this 'microfunctional' solution is open to scientific functionalists, but not to common-sense functionalists. Assumptions about reflectance profiles and neuronal firing are not part of common-sense psychology. Common-sense functionalism is therefore forced to count Matthew as functionally normal, and so unable to explain his intuitive qualitative difference by reference to functional factors. Scientific functionalism is not so constrained, as there are many possible levels of functional organization in the brain which might distinguish Matthew from normal humans. The only question is whether this is an embarrassment of riches. For once we have opened up the possibility that qualitative features might be fixed by any level of functional organization, what could possibly decide which level does in fact fix it? It is no good appealing to the subject's introspective reports, since they are part of macro-functional role, which is already agreed not to fix qualia (thus Matthew will describe state A as 'seeing red', yet nearly everybody agrees he is experiencing green). On the other hand, it seems highly unlikely that intuition can resolve the issue, for intuition is surely not fine-grained enough to decide such questions as whether somebody who is computing the normal reflectance profile for green things, but using an abnormal pattern of neuronal firing to do so, is experiencing green or not.

6. Absent qualia

The inverted spectrum thought experiment is consistent with the idea that appropriate functional organization guarantees conscious states of some kind. The only issue raised by the inverted spectrum argument is whether functional identity suffices to fix the qualitative identity of those conscious states. A more radical form of argument asks whether functional organization can guarantee consciousness at all. Are there not always going to be systems that display the appropriate functional organization yet lack qualia altogether?

Clearly, answers to this question will depend crucially on what qualifies as 'appropriate functional organization'. Recall that earlier the notion of a functional role was introduced via Ramsey's strategy for defining unobservable terms in terms of a theory. On this conception,

you have mental states with a certain functional role if you have a set of states which interact in the way specified by some psychological theory $T(M_1 \dots M_n, O)$.

A number of thought experiments aim to show that any such notion of functional role must yield too 'liberal' an account of qualitative states, in that it will imply the presence of qualia in systems that in fact lack them. Imagine that the sensory messages into your brain and the motor messages going out are disconnected from your brain and instead linked by radio to a complicated structure made of old beer cans. This system of beer cans is arranged in such a way that its properties relate to each other and to your sensory and motor messages in just the way that psychological theory T says your brain states did before they were disconnected. Then the system consisting of your debrained body plus beer cans will have exactly the same functional organization as you had before your brain was disconnected. Yet it seems implausible that this system would be conscious, that there is something that it would be like to be a debrained body plus beer cans.

Even so, a significant number of functionalist philosophers are prepared to bite this bullet, and maintain that the beer can system would be conscious. After all, if some diseased part of your brain were replaced by some microscopic silicon-based circuitry, this would not necessarily stop you being conscious. What is the difference in principle, it could be asked, between this and the beer can case?

The beer can thought experiment starts with a normal human being, and then assumes the alteration of this human being's brain. But there are further arguments which query whether even behavioural similarity to human beings is necessary for the relevant kind of functional organization. These arguments focus on the contribution of the observational terms O to the specification of functional roles. Note that it is essential to any plausible functionalism that these terms somehow be independently understood. For, if the O terms are themselves defined in terms of theory T , there will be far too many systems the properties of which satisfy T . Indeed, given a completely uninterpreted theory $T(M_1 \dots M_n, O)$, it is arguable that *any* physical system – the air molecules in your house, say – will display *some* set of properties that are related to each other as T says the M s and O are related. So some prior understanding of the O terms is needed to give us any chance of ensuring that only conscious systems have the functional organization specified by T .

The natural way to do this is to equate O terms with terms for perceptual inputs (for example, 'receiving visual stimuli from an ice cream') and behavioural outputs (for example 'moving your arm towards an ice cream'). But this, then, has the disadvantage of implying that some presumably conscious human beings, such as experimental subjects whose optic nerves are being stimulated by brain scientists, or paralyzed people who display no behavioural outputs, will lack the relevant functional organization.

The natural response to this difficulty (for a scientific functionalist at least, even if not for a common-sense functionalist) is to appeal to physiology to specify less peripheral inputs (kinds of stimulation of the optic nerve, say) and less peripheral outputs (kinds of activity in the motor cortex). But this physiological move threatens chauvinism. For it will exclude from the category of conscious beings any extraterrestrials and animals whose non-human physiologies do not include optic nerves and motor cortexes.

There remains some room for functionalists to manoeuvre here. One possibility would be to return to Lewis' mixed theory, which would imply that the experimental subject and the paralyzed person are conscious because they are in the states that play the requisite causal role in their normal conspecifics, even if not in themselves. Another increasingly popular response is to appeal to teleological considerations, and understand 'functional organization' as not only specifying a causal structure but requiring in addition that this structure be a product of biological design. By appealing to this further sense of 'function', such 'teleofunctionalists' can specify that beings with different physiologies can nevertheless share the same functional organization, in virtue of having causally similar structures which have been designed for similar biological purposes.

7. Representational content

This further teleological sense of 'function' is also relevant to the difficulties functionalists face in accounting for content. Many mental states, and in particular such non-qualitative propositional attitudes as belief and desire, represent states of affairs other than themselves. It is unclear whether the functionalist theory of mental states can account for such representational contents.

One immediate difficulty is that many representational contents are 'broad', in the sense that the content of many propositional attitudes seems to depend not just on the thinker's physical make-up but also on features of the context (see Content: wide and narrow). For example, it is arguable that the possession of beliefs about natural kinds depends not just on the organization of the believer's brain, but also on features of the believer's social context and on which natural kinds are present in the believer's world. This raises a *prima facie* problem for functionalism. For functionalism makes the possession of mental states depend on the subject's brain having a certain causal structure. Yet by hypothesis the possession of broad propositional attitudes is not fixed by facts about brains.

However, functionalism has an obvious remedy. There is no obvious reason why functionalism should only consider causal structures inside the head. Why not have a 'broad' functionalism, which recognizes 'broad' causal structures in which mental states interact, not only with each other and with sensory inputs and motor outputs but also with such external factors as the social

environment and objectively existing natural kinds? This would then open the way to a theory which makes the possession of broad propositional attitudes depend on broad causal structures. So perhaps broad contents present no special problem for functionalism. A more radical criticism of functionalism, however, queries its ability to account for representational contents of any kind, broad or not. Functionalism identifies mental states as items that have certain causes and effects. Yet it is doubtful whether representation can be explained in any simple causal terms.

It might seem that, once we are allowed to appeal to 'broad' causal structures, we can identify the contents of beliefs as those external circumstances that typically cause them, and the contents of desires as those external states of affairs to which they typically give rise. But this simple strategy is afflicted by the problem of 'disjunctivitis'. Take the belief with the content that an ice cream is in front of you. This can be caused, not only by a real ice cream, but also by a plastic ice cream, or a hologram of an ice cream, and so on. Similarly the results which follow any given desire will include not only the real content of the desire, but also various unintended consequences.

So, even if we are allowed broad causal roles that include external causes and effects, we still need somehow to identify, among the various causes that give rise to beliefs, and the various results that eventuate from desires, those which comprise the beliefs' and desires' real contents. There are a number of possible ways of doing this. One of the most promising is to appeal to teleological considerations once more. For then we can pick out a desire's content as that effect which it is the desire's biological purpose to produce. And, similarly, we can pick out a belief's content as that condition with which it is the biological purpose of the belief to be co-present (see Semantics, teleological).

One question raised by this appeal to teleology is whether the original functionalism is still doing any work. We started with the functionalist idea that mental states can be identified by their causal roles. But it now seems that, for contentful mental states at least, causal roles are not enough, and need to be supplemented by biological purposes. The obvious question is whether biological purposes would suffice by themselves. Perhaps we can identify contentful mental states by their purposes alone. The answer depends on whether a common biological purpose, but different causal roles, implies different representational states, and relates to a number of currently controversial questions in the philosophy of representation.

References and further reading

Block, N. (1980) Readings in the Philosophy of Psychology, London: Methuen, vol. 1. (The section on functionalism contains many important articles, including early versions of functionalism by H. Putnam, D. Lewis and D.M. Armstrong. See also Block's Introduction to this section, and the difficulties he raises in 'Troubles with Functionalism'.)

Descartes, R. (1641) *Meditations on First Philosophy* in *The Philosophical Writings of René Descartes*, trans. J. Cottingham, R. Stoothoff, D. Murdoch, Cambridge: Cambridge University Press, vol. 2, 1984.

(Classic statement and defence of dualism.)

Lewis, D. (1980) 'Mad Pain and Martian Pain', in N. Block (ed.), *Readings in the Philosophy of Psychology*, London: Methuen, 1980, vol. 1; repr. with a postscript in *D. Lewis Philosophical Papers*, Oxford: Blackwell 1983, vol. 1.

(Lewis' 'mixed' version of functionalism.)

Millikan, R. (1986) 'Thoughts without Laws', *Philosophical Review* XCV: 47–80.

(Argues that mental states should be analysed in terms of biological purposes rather than causal roles.)

Nagel, T. (1974) 'What is it like to be a Bat?', *Philosophical Review* 83: 435–450, repr. in Block (1980).

(Argues that the subjective character of experience is not captured by physicalist theories.)

Papineau, D. (1991) 'Teleology and Mental States', *Aristotelian Society Supplementary Volume* LXV: 33–54.

(Argues that functionalism without teleology cannot account for representation.)

Ramsey, F.P. (1931) 'Theories' in *Foundations of Mathematics*, London: Routledge & Kegan Paul.

(The original statement of Ramsey's account of theoretical terms.)

Rosenthal, D.M. (1991) *The Nature of Mind*, London: Oxford University Press.

(Good general collection on contemporary philosophy of mind. Papers by W.G. Lycan and E. Sober argue for 'microfunctionalism' (Lycan) and 'teleofunctionalism' (Lycan and Sober).)

Shoemaker, S. (1984) *Identity, Cause and Mind*, Cambridge: Cambridge University Press.

(This collection of Shoemaker's papers includes discussions of the different species of functionalism and of the inverted spectrum and absent qualia objections.)