Mind Association

http://www.jstor.org

Functionalism and Broad Content

FRANK JACKSON AND PHILIP PETTIT

The contents of certain intentional states are broad or context-bound. The contents of some beliefs (and desires and hopes . . .) depend on how things are outside the subject in addition to depending on how things are inside the subject.[1] What implications does this have for the functionalist theory of mind? In this paper we defend the simplest reply to this question.

Functionalism is the doctrine that, for very many kinds of psychological states, to be in a psychological state of that kind is to have in one a state playing a certain role between inputs, outputs, and other internal states. It is the nature of the role, not the nature of the occupant of the role, which matters. The simplest functionalist response to states with broad contents is to analyse them in terms of broad roles, that is in terms of roles which are specified as having some inputs or outputs that are happenings outside the skin. When functionalists give an account of a belief with broad content, this response requires them to include outside happenings in their specification of the functional role definitive of having a belief with that content. The response, therefore, ensures that having a belief with that content involves how things are outside the subject as well as how they are inside.[2]

It has seemed to many that this simple, minimally disruptive response on behalf of functionalism to the fact of broad content cannot be right. Sometimes the argument turns on the claim that because an intentional state's content is an essential property of it, states with broad content are essentially broad, and essential broadness is alleged to be incompatible with functionalism, or at least incompatible with the functionalism just sketched. The idea is that broad content presents a much more radical challenge to

functionalism's whole perspective on the mind than our tinkering acknow-
ledges. As it is sometimes put, broad content shows the failure of the view
that mental processes are essentially inner processes, and it is alleged that
functionalism is wedded to the inner picture.[3]

Other times the objection comes from the functionalist camp itself. It is
claimed that a certain argument, which we will call the doppelgänger
challenge, shows that the ascription of broad content does not play an
appropriate role in the explanation of behaviour. It is argued that it is the
ascription of narrow content, content which does not depend on how things
are outside the subject, which plays the appropriate role in explaining
behaviour.[4] Accordingly, we should be dualists about content. There is
narrow content, which plays the right kind of explanatory role with respect
to behaviour to fall within the purview of functionalism, and there is broad
content, which plays a distinct role, perhaps that of representing the world,
or perhaps that of being denigrated as a relic of folk psychology. The most
plausible version of the dualist response casts states with broad content as
mixed states. They are in part properly psychological—this part has narrow
content and is given a functionalist account; and they are in part not
psychological at all—this part is constituted by the relations between a
subject and his environment and is no business of psychology, and so no
business of functionalism.[5]

Call a functionalism which insists that the inputs or outputs in its
functional roles sometimes include external happenings 'broad', and one
that confines the inputs and outputs to happenings at or inside the skin
'narrow'. Our position is that functionalists should respond to broad
contents by moving to broad functionalism, analysing such contents by
broad roles. The first objection is then that broad functionalism cannot
accommodate the fact that broad content is essential, and the second
objection is that broad functionalism cannot accommodate the fact that, as it
is argued, broad content is not suitably explanatory.

We reply to the first objection in part I where we argue that broad
functionalism is neutral on the question of whether broad content is
essential or accidental. In part II we reply to the second objection by
arguing that the ascription of broad content is explanatory of behaviour in a
way fully in accord with the motivation of functionalism and with a properly
scientific attitude towards psychology. The bulk of this part is concerned
with what we will call the doppelgänger challenge to broad content's
explanatory value.

[3] See, e.g., John McDowell, 'Singular Thought and Inner Space', in Pettit and McDowell, op. cit.

[4] For a recent forceful presentation, see Jerry Fodor, 'Individualism and Supervenience I',
Proceedings of the Aristotelian Society, Supp. Vol. LX, 1986, pp. 235-62. We are using 'narrow
content' here in a catch-all way to include for instance John Perry's senses, see n. 11.

[5] See Fodor, op. cit., and, for a particularly detailed account of a dualist view, Colin McGinn, 'The
Structure of Content', in Andrew Woodfield (ed.), *Thought and Object*, Oxford University Press, 1982.
For further dualist accounts see the references in Owens, op. cit.

I. FUNCTIONALISM AND ESSENTIAL BROADNESS

Why essentially broad content is thought a problem

Broad content is content which does not supervene on how things are inside me. When I am in an intentional state with broad content, things must be thus and so around me as well as being thus and so inside me. For later purposes, particularly in part II, it is helpful to put the distinction between narrow and broad content in terms of the notion of a doppelgänger. A doppelgänger of mine is a subject of experience exactly like me as far as what goes on inside the skin is concerned. He is an internal molecule for internal molecule duplicate of me. Narrow content is content subjects share with all their doppelgängers; broad content is content subjects do not share with all their doppelgängers.

(There is a certain degree of arbitrariness about drawing the boundary at the skin. Why not count a brain in a vat which duplicates everything that goes on in my brain while differing profoundly in respect of what goes on between my brain and my skin, as one of my doppelgängers? It is, however, a little easier to conduct the discussion in terms of the narrower definition, for it means that the similarities in outputs between me and my doppelgängers can be described (roughly, anyway) in familiar behaviourial terms rather than in neurophysiological ones.)

We are going to take for granted the familiar and persuasive arguments for the existence of broad content; and, in order to focus the discussion, we are going to take as our standard illustration of an intentional state with broad content my belief that that cup contains water. We will suppose that the content which the sentence 'that cup contains water' assigns to my belief is such that two conditions need to be fulfilled. First, I could not have a belief with that content unless it was that very cup which I was causally affected by—say, by seeing it. Thus, my doppelgänger viewing an exactly similar but numerically distinct cup, has a belief with a different content (whether he knows it or not). And, secondly, I could not have a belief with that content unless I had causal commerce, even if only through a baptismal chain, with water. This is why my twin earth doppelgänger does not have a belief with that content (he may use the same word 'water' in characterizing his belief, but the content he ascribes will nevertheless be different). This kind of position on intentional states like my belief that that cup contains water is very widely accepted, and general agreement with a position of this overall kind is all that we need for the paper.[6]

Many hold that a psychological state's content—be it broad or narrow—is an essential property of it. Consider our sample state, my belief that that cup contains water. Had it been a different cup or had I been on twin earth, I

---

[6] See, e.g., the papers cited in n. 1.

could not have been in that state, not just in the sense that the description 'is a belief that that cup contains water' would not have been true of my state—that is, not just in the sense that my state might have had a different content—but in the sense that I could not have been in that very state, however described. This view has the consequence that some mental states, in particular those with broad content, are essentially broad, and it lies at the heart of much of the controversy over broad psychology. For it is the view that had the world around me been different in various ways, I would *have* to have been in different mental states—not merely in the sense of states satisfying different descriptions but in the sense of numerically different states. It is this view which challenges the 'inner life' view of psychology, for it makes the very identity of my mental states a matter of my surroundings.[7]

A natural thought is that broad functionalism is incompatible with the existence of essentially broad states. If there are essentially broad psychological states, then functionalism (broad or narrow) is false, or at least false as a theory of those states. We will argue, on the contrary, that functionalism *per se* is neutral on the existence of essentially broad states. It is functionalism as it is usually characterized which is incompatible with the essentially broad, and the relevant part of this characterization is entirely optional for functionalists.

Two possible metaphysics for functionalists

Functionalism *per se* is a thesis about the truth-conditions of psychological attributions. It says that I am in mental state M when I have in me a state which is playing the causally intermediate role between inputs, outputs, and other states definitive of M. For short, I am in M if and only if I have in me a state playing the M role. The metaphysical question of which state *is* the psychological state is a further matter. Thus functionalism *per se* says that to be in pain is to have in one a state playing the pain role, where the pain role is defined in terms of being causally intermediate between inputs, outputs, and other states in one. And this means that there are *two* candidates to be one's pain: the state which plays the role, and the state of having the state which plays the role. Suppose that it is having C fibres firing which plays the pain role in me on some occasion, then the essence of the functionalist view is that I am in pain precisely because this state is playing that role in me. This leaves open which of two states are my pain: my having C fibres firing or my having in me the state—which is, as it happens, C fibres firing— which plays the pain role.

Consider the following parallel. To be sky-coloured is to have the same colour as the sky. That gives us the conditions under which 'My house is sky-coloured' is true, namely, that it is true if and only if 'My house is the

[7] See, e.g., McDowell, 'Singular Thought and Inner Space', op. cit., and Tyler Burge, 'Individualism and the Mental', op. cit., and 'Other Bodies' in Woodfield, op. cit.

same colour as the sky' is true. Suppose that my house is in fact the same colour as the sky, which state of my house *is* its state of being sky-coloured? Its state of being blue, or its state of being the same in colour as the sky? There are two states here, for had the sky and my house both been red, my house would not have been in the first state but would still have been in the second state. But although there are two states, it does not follow that one answer or the other is definitely correct as to which state is my house's being sky-coloured. Perhaps, 'my house being sky-coloured' is variable and ambiguous in its reference: in some contexts it refers to one state, in some contexts to the other, and in still other contexts there is no fact of the matter as to which of the two states it refers to. For suppose someone expresses surprise at my house being sky-coloured, what are they expressing surprise at? The fact that my house is blue, or the fact that my house is the same colour as the sky? Clearly there are two distinct possibilities here, but stipulation as to the intended reference of 'my house being sky-coloured' is called for in order to know which possibility is being remarked upon.

The truth-conditional essence of functionalism leaves open the metaphysical question of which state—my having my C fibres firing, or my having the state playing the pain role—is my state of pain. The usual option taken is to make my having C fibres firing be my pain, and in general to make the state playing, realizing, or occupying the role, call it 'the realizer state', be the psychological state; the less usual option is to make my having the state playing the pain role be my pain, and in general to make what we will call 'the role state', the *having* of a state occupying the role, be the psychological state. The role state is a second-order state, namely, that of having a state of a certain kind, just as having the same colour as the sky is a second-order property. The distinction between the role state and the realizer state is a subtle one, and surely it is plausible—from a functionalist perspective anyway—that the natural language terms we use for denoting the various psychological states are ambiguous and variable as regards which of the two is being referred to. A decision between the two is a matter of stipulation, not discovery. In any case, whatever should be said about linguistic usage, functionalists may properly count either role states or realizer states as being what psychological states are. There is nothing in functionalism *per se* which rules out either option.

Reply to the objection from essential broadness

If we take the usual option of identifying psychological states with realizer states, it is quite correct that functionalism is incompatible with essential broadness, though it easily and naturally accommodates what we might call accidental broadness. Take any mental state with a broad content. On the broad functionalist account, the broadness of that state is captured simply by making the inputs or outputs associated with its role include happenings

outside the skin. For instance, the functionalist story about my belief that that cup contains water will include the proviso that the state be caused (in the right way) by that cup and by water. But, on the realizer option for functionalism, the broadness secured by including this proviso will be an accidental broadness. For the very same state might of course have stood in different causal relations: the functional role a state plays is not an essential property of it. Thus, the broad functionalist who takes the realizer state option must say that my belief that that cup contains water, that very state, might not have been caused by that cup, but by, say, a duplicate of that cup, and might not have had the causal links to water, that is, to H_2O, which are responsible for it being, on the account in question, the belief that it is water and not XYZ which the cup contains.

Suppose, however, that a broad functionalist takes the less usual option of identifying my belief that that cup contains water with the role state instead of with the realizer state. My belief then will be, not the state which plays the belief-that-that-cup-contains-water role, but the state of having a state which plays the belief-that-that-cup-contains-water role. Now for this second-order state, though not for the realizer state, individuation goes by role: different role, numerically different state. Although the very same realizer state might have been caused by this cup or instead by that one, or by XYZ instead of by H_2O, my being in a realizer state caused by this cup rather than by that one or in a realizer state caused by H_2O rather than by XYZ are different role states. Replace water then by XYZ, or that cup by another, no matter how similar, and it will no longer be true that I believe that that cup contains water. Why? Because the functional role realized is different; but that is to say that I am in a numerically different role state, though I may be in numerically the same realizer state. The upshot is that, compatibly with functionalism, we may take my belief that that cup contains water to be essentially broad, for we may take my belief to be the role state, and the role state's identity—which state it *is*, not just which descriptions it satisfies—depends on how things are outside of me, in particular, on whether or not it is water and that cup which do the causing of the realizer state.[8]

On two objections to our metaphysical neutrality

Our claim is that functionalists, and in particular broad functionalists, have a choice between identifying mental states with realizer states or with role states, and that therefore broad functionalism is neutral as between whether

---

[8] Our argument that broad functionalism can allow essentially broad states rests on the claim that role states are individuated in part by role. If you are unhappy about this claim—perhaps because it is token states which we are talking about, and you take a coarse-grained view on the individuation of token states, even when they are second-order states—replace role states in the above by aggregations of realizer states with the relevant external happenings. As aggregations with different members are necessarily distinct, this makes broadness an essential feature.

my belief that that cup contains water is accidentally or essentially broad. There is, however, one consideration which suggests that functionalists ought to take the realizer state option, and another which suggests that they ought to take the role state option. If either is correct, our neutrality is misconceived.

If we want psychological states to be causes of behaviour, then we must regard them as realizer states. It is the state that fills the functional role, not the state of having a state that fills the role, which does the causing of behaviour. And this looks like a powerful consideration in favour of identifying mental states with realizer states. Was not the explanatory role of the mental a major reason for moving from behaviourism to functionalism?

What, exactly, is the datum concerning psychological states and causality? Perhaps in the case of the bodily sensations there is introspective evidence that they themselves cause behaviour, though even there this sort of evidence should be treated with extreme circumspection. In the case of beliefs and other intentional states, however, the only clear datum is that we invoke them in causal explanations of behaviour, not that they are causes of behaviour; beliefs are not inner feelings whose causal links are available to introspection. Now the role state conception of belief is perfectly compatible with, indeed entails, that ascribing a belief of a certain kind explains behaviour, for on that conception to ascribe a belief is to say that there is a cause of a certain kind in one who has the belief, even though the belief is not itself the cause. I may explain why an object moves in a certain way by saying that there is a force of a certain kind acting on it, even though it is the force itself, not the existential state of affairs of there being a force of the certain kind, which does the causing.

The consideration that appears to favour the role state option relates to a point made at the very beginning. We observed that if, as many hold, an intentional state's content is an essential property of it, then to admit broad content is to admit that some states are essentially broad. Thus Tyler Burge argues that we individuate beliefs by content, so that different content means, not just different description of belief state, but numerically different belief state.[9] Suppose that the links to that cup and to water are broken, so that the content of the belief changes and the belief I have is no longer correctly described as the belief that that cup contains water. Burge would say that to suggest that we might still be dealing with the same state is as absurd as the suggestion that by changing the membership of a set we might change the descriptions true of the set but not change the set that is being described. If Burge is right, functionalists cannot take the realizer state option, at least for states with functionally specified content, for that option makes content a contingent feature.

The appeal of this argument rests on conflating the two senses in which

[9] See the papers by Burge cited in n. 7. The view that content is essential cuts across many differences in doctrine elsewhere, thus Fodor, op. cit., endorses it.

we talk about beliefs, thoughts, and the propositional attitudes in general. It is often observed—for instance in discussing the bad objection to the identity theory that thoughts are not located whereas brain states are—that we sometimes use the term 'thought' to denote what is thought, the content, and sometimes the thought itself, the state located in space-time. The claim that beliefs are individuated by content is beyond question if meant as a claim about what is thought; indeed, it is then a tautology. But this claim is very much in question if it is meant as a claim about the state of belief, the state with the content. Functionalists of the realizer state persuasion are entitled to reply that they have provided a plausible account of what a belief is, and that this account has the consequence that content does not individuate: that is, that a belief's content is not an essential property of it. This reply turns on the point that which properties of a thing are essential may be an a posteriori question. But this is hardly controversial: orphans do not refute the necessity of origin.

We hope to have defused the issue about the compatibility of essentially broad mental states with broad functionalism. Essential broadness is a free option for broad functionalists. To take it, formulate your functionalism with role states being the psychological states. To decline it, formulate your functionalism in the more usual way so that realizer states are the psychological states. The states with broad content will then be accidentally broad and you need a reply (we have suggested one immediately above) to those who hold that content is essential.

II. THE EXPLANATORY VALUE OF BROAD CONTENT

The ascription of content to an intentional state is part of the project of explaining a subject's behaviour. The broad functionalist accepts this, adding that the point applies as much to broad as to narrow content. Given that I desire water, my belief that that cup contains water explains why I reach for that cup in the same general way that my belief that the only white cup in the room contains stuff that goes under the name of 'water'—supposing for the purposes of exposition that the latter is indeed a belief with narrow content—explains why I reach for the only white cup in the room. In this part we reply to the familiar challenge, the doppelgänger challenge, to the explanatory value of broad content. We start with some clarificatory remarks about the nature of the challenge before explaining the notion central to our reply to it, that of a distinction between those causal explanations we will call programme explanations and those we will call process explanations.

We will frame our discussion in terms of a functionalism which identifies intentional states with realizer states, this being the more familiar way of thinking of functionalism. However nothing we say about the explanatory

value of broad content turns on this choice—as had better be the case given our relaxed stance on the decision between the role and realizer state options for functionalism. The metaphysics of broad content is a separate matter from the explanatory value of broad content. Our question is whether the ascription of certain connections between an agent's internal goings on and external happenings can be explanatory of his behaviour in the way required by the functionalist perspective. The nature of this question is not affected by whether we hold that those goings on are the states with broad content, or whether it is the second-order states of having states with such connections which are the states with broad content. Either way the internal states are there and are so connected with external affairs, and it is the explanatory significance of this fact which concerns us.

The doppelgänger challenge

The doppelgänger challenge to the explanatory value of broad content is sometimes put rather tersely, somewhat as follows.[10] I and my doppel- gängers behave exactly alike. That is, we agree in behaviour. Also we agree in narrow states while disagreeing in broad states. Mill's method of agreement leads immediately to the conclusion that my narrow states explain my behaviour, and that my broad states are explanatorily idle. (And the same goes, of course, for each of my doppelgängers.)

This is wrong in two ways. First, I and my doppelgängers do not behave exactly alike. I reach for that cup and drink water, my doppelgänger on twin earth reaches for a different cup and drinks XYZ. Secondly, we have here a highly dubious application of the method of agreement. Suppose my pulling the trigger of a gun explains someone's death. Consider the bullet as it enters the body and imagine it replaced by a doppelgänger bullet moving with the same velocity, entering at the same point, at the same angle, and so on, but issuing from another source: your gun, not mine. The death will be exactly the same in every detail, yet this in no way shows that my pulling the trigger was not part of the explanation of the death. I cannot get off the hook with a thought experiment! The reason the argument goes wrong here is that my pulling the trigger is part of the explanation of why the bullet was as it was when it entered the body. Similarly, broad content and narrow content are not independent in the way required for the method of agreement to show that broad content is idle. What makes it true that I am in a state with broad content is in part my causal past, and that causal past will be part of the explanation of my being in a state with a certain narrow content. What that cup and various samples of water did to me in the past is in part responsible for those narrow states which lead me to reach for that cup.

The substantial doppelgänger challenge, the one that calls for a detailed reply in terms, as we will argue, of the notion of a programme explanation is

10 See, e.g., Michael Devitt and Kim Sterelny, *Language and Reality*, Basil Blackwell, 1987, p. 166.

not designed to show that broad content is explanatorily idle, but rather that it is explanatorily redundant: that is, that it plays no distinctive role, the explaining it does being best done by a combination of narrow psychological states together with environment.[11] Although my behaviour is not exactly the same as that of my doppelgängers, it is similar in an important respect. When I reach for that cup and drink the water in it, many of my doppelgängers will be reaching for a different cup and drinking XYZ. This is different behaviour, but obviously not that different. We all reach out and we all swallow, while differing in what we reach out towards and in what we swallow. There is, that is, a set of core behaviours in common between me and all my doppelgängers. Our bodies gyrate in the same way, as we might put it. How is this fact to be explained? Obviously, in terms of the fact that I and my doppelgängers are similar in a very important way; but that way does not include sharing psychological states with broad content. It is only narrow psychological states that we share. Hence, runs the substantial doppelgänger challenge, as far as explanation in psychological terms goes, we only need narrow content to explain that aspect of my behaviour which I share with my doppelgängers.

If we call this shared aspect of behaviour, pure behaviour, the point is that narrow content is all we need to explain pure behaviour. How then is impure behaviour to be explained, how is my reaching for *that cup* or my swallowing *water* to be explained? The answer seems clear. What needs to be added to swallowing to get swallowing water is water, what needs to be added to reaching to get reaching for that cup is that cup. In general, then, the suggestion is that we can explain impure behaviour by adding to our explanation of pure behaviour facts about the environment. But then we can explain behaviour, be it pure or impure (or, perhaps better, be it described in pure or impure terms) in terms of psychological states with narrow content together with, if needed, facts about the subject's environment. The conclusion is that broad content is redundant, that it plays no distinctive role in explaining behaviour.

We can reinforce the point that the substantial doppelgänger challenge is not that broad content plays no explanatory role at all, only that it plays no distinctive one. The line of thought which underlies the doppelgänger challenge naturally suggests a picture which *entails* that broad content explains behaviour. Suppose that, on some given occasion, I believe that that cup contains water. And suppose we accept, as we did in expounding the challenge, that though my doppelgängers do not share this belief, there is some mental state which we do all share, call it the belief that < that cup contains water >, which together with the appropriate environmental details entails that I believe that that cup contains water. That is, I both

---

[11] Although some of the phrasings in Colin McGinn, 'The Structure of Content', in Woodfield, op. cit., suggest that he holds that broad content is idle, we take his considered view to be that it is redundant.

believe that < that cup contains water > and that that cup contains water, and what makes the former up into the latter is my environment.[12]

It may be that we did not need to introduce the special ' <   > ' notation here and that ordinary English will do. Perhaps the underlying narrow belief of mine is that there is a white cup, and only one white cup, which contains stuff known as 'water' in my speech community. Or perhaps not even this belief is narrow in content, and we need the language of thought instead of English. Or perhaps it is not really a belief, but a proto thought or belief, which is the psychological state underlying my belief that that cup contains water.[13] None of this matters here. The crucial point is that on this natural way of fleshing out the doppelgänger challenge, it follows from that challenge that broad content has explanatory value. For to ascribe broad content is to commit oneself to there being a psychological state with narrow content which, according to the challenge itself, is what does the explaining. Hence, by this very fact, the ascription of broad content is explanatory. If to ascribe broad content is *inter alia* to ascribe narrow content and narrow content is explanatory of behaviour, then so is broad content.

The doppelgänger challenge, then, should be seen as seeking to induce, not scepticism about the explanatory value of broad content, but scepticism about the *distinctive* explanatory value of broad content. What explanatory work does it do that is not done by the correlative narrow content (however exactly we pick out that narrow content)? We now seek to answer this question, and so the doppelgänger challenge, via the idea of a programme explanation.

Programme explanations

The doppelgänger challenge is only one way of raising the problem of how the ascription of broad content explains. Sometimes the problem is raised by drawing on a distinction between those properties which are causally efficacious or productive and those which do nothing. (Sometimes the term 'causally explanatory' is used in place of 'causally efficacious', but as we will see shortly this obscures a crucial distinction.) It is observed that this coin landing heads rather than tails does not affect the rising of the sun tomorrow. True, it changes that event in the sense that had the coin not landed heads, the rising of the sun would not have had the property of occurring after the coin landed heads, but that is a 'Cambridge' change in the rising of the sun. The fact remains that we have here a good example of

---

[12] In the terminology of John Perry, 'Frege on Demonstratives', *Philosophical Review*, 1977, pp. 474-97, the *sense* of what I believe is common to me and my döppelgangers, though we have different *thoughts*. Perry argues that it is the sense of what I believe and not the thought, that is important for psychological explanation.

[13] For a view of this kind see Daniel Dennett, 'Beyond Belief', in Woodfield, op. cit.

an inefficacious property, however difficult it may be to give a philosophically precise account of the notion.[14]

Similarly, what makes a psychological state broad is a good example of a causally inefficacious property, for what makes it broad is irrelevant to its effect on other states, and so to its effect on behaviour. What makes a state my belief that that cup contains water is, in part, its causal connection with that cup and with water. We include that cup and water among the inputs in the broad functionalist story, as opposed, for instance, to including a duplicate of that cup and XYZ. Now these causal connections will be quite invisible to the other states of mine which link my belief to behaviour. For example, if my belief is my having B-fibres firing, then the effect of my belief on other states of mine along the causal path to behaviour will simply be the effect of B-fibres' firing. The fact that the B-fibres' firing was caused in a certain way will be quite invisible to the other states.

Raised in this way, the problem concerns narrow functionalism as much as broad functionalism. The argument turns on the highly relational nature of broad content, not on its non-individualistic nature, and narrow content, though individualistic, is highly relational. For the effect that one neurophysiological state has on another in the causal path to behaviour is a function of the states' relatively intrinsic properties, not of their highly relational ones. What one neurophysiological state does to another neurophysiological state is a function of their relatively intrinsic characters, not of their relations to more remote states, regardless of whether those more remote states are inside or outside the skin. How then can the possession of a certain narrow content by an agent's intentional state—a highly relational matter according to functionalism—explain the agent's behaviour?

Our argument will be that this sort of argument, be it deployed against broad or narrow content, is misconceived, and that seeing why it is misconceived shows the mistake in the doppelgänger challenge. The argument embodies a plausible but quite mistaken view about causal explanation, namely that causal explanations that proceed by citing some feature must cite a causally efficacious or productive feature. On the contrary, we can and often do explain by citing a feature which causally *programmes* without causing. Features which causally explain need not cause. This is typically what happens when we explain in terms of highly relational properties. We will introduce the idea of causally programming through some everyday examples of explanations.

Electrons A and B are acted on by independent forces F_A and F_B, respectively, and electron A then accelerates at the same rate as electron B. The explanation of this fact is that the magnitude of the two forces is the same. This is a perfectly acceptable explanation, and yet the same points that are made against broad psychological explanation can be made about it.

[14] Cf. Fodor, op. cit., p. 243.

We explain the fact that electron A accelerates at the same rate as B in terms of the force acting on A being the same in magnitude as that acting on B. But this sameness in magnitude is quite invisible to A and is not shared by A's doppelgängers. This sameness does not make A move off more or less briskly; what determines the rate at which A accelerates is the magnitude of F_A, not that magnitude's relationship to another force altogether. The two electron-force pairs may indeed be sufficiently separated to be quite independent of each other. What happens to one does not causally affect what happens to the other. Or, in other words, the equality *per se* of the forces acting on the electrons does not do any causal work. The work is all done by the individual forces acting on the electrons. (And if the equality *per se* did do some work, how remarkable that it arranges to do exactly the right amount so as not to conflict with the result produced by the individual forces.)

There is nothing particularly special about the electron example. It is typical of many cases where a present fact is explained in terms of a highly relational past fact. 'Why was the price of a lamb at the Portland auction the same as the price at the Ballarat auction? Because the relationship between supply and demand was the same at both places.' 'Why did the Bourbons last longer than the Windsors? Because they had more male offspring.' 'Why do trees grow faster in Melbourne than in Canberra? Because there are fewer frosts in Melbourne.' In each of these cases, remarks similar to those we made about the electron case apply. For example it is the number of frosts that slows down the trees in Canberra, not the fact that there are fewer than in Melbourne. The trees in Canberra would grow at exactly the same rate whether or not there was a single frost in Melbourne.

How is it that in these examples the instantiation of a property causally explains a certain result without actually producing it? The answer is that we have in these examples a range of possible situations, any one of which would be equally able to produce the result, and one of which actually does produce the result. The impressed forces on the electrons are, say, each of magnitude five, causing an acceleration in each electron of magnitude ten, say. Thus, in this case what actually produces the result that the accelerations are the same is both forces being of magnitude five. But both forces being of magnitude six, or seven, or . . . instead would equally have produced the result that the electrons' accelerations were the same. The property we cite as explaining the result is the relevant property in common between the various members of the range of possible situations, each member of which would have produced the result and one of which did in fact produce the result. We will describe such a common property as causally *programming* the result, and call explanations in terms of properties which programme without producing, *programme explanations*, as opposed to *process explanations* which do cite the productive features.

We can express the basic idea behind a programme explanation in terms

of what remains constant under variation. Suppose state A caused state B. Variations on A, say, A′, A″, . . . would have caused variations on B, say B′, B″, . . ., respectively. It may be that if the A^i share a property P, the B^i would share a property Q: keep P constant among the actual and possible causes, and Q remains constant among the actual and possible effects. If you like, Q *tracks* P.[15] Our point is that in such a case P causally explains Q by programming it, even though it may be that P does not produce Q.[16]

Explanations in terms of highly disjunctive states are typically programme explanations. It is not the disjunctive state *per se* which produces the result but rather the state corresponding to one or other of the disjuncts. But we accept the explanation stated in the disjunctive terms provided any state, or any state within limits, corresponding to one of the disjuncts would have produced the result. We may explain the conductor's annoyance at a concert by the fact that someone coughed. What will have actually caused the conductor's annoyance will be the coughing of some particular person, Fred, say; but when we say that it was someone's coughing that explains why the conductor was annoyed, we are thinking of someone's coughing as Fred's coughing or Mary's coughing or Harry's coughing or . . ., and saying that any one of these disjuncts would have caused the conductor's annoyance—it did not have to be Fred.

Explanation by properties which causally programme rather than produce what is to be explained is very common. Take Hilary Putnam's famous example where the failure of a one inch peg to go into a one inch round hole is explained by the fact that the peg is square. As Putnam observes, the causal interactions will all be between particular sub-microscopic particles, and that story will not involve squareness as such. The squareness of the peg and the roundness of the hole will enter the picture only indirectly, namely by virtue of the fact that there is a whole range of possible situations constituting a square one inch peg interacting with a round one inch hole; in all of them (barring the extremely unlikely) the peg will not fit into the hole. As we say it, the squareness of the peg together with the roundness of the hole programmes the failure of the peg to fit into the hole, or the failure of the peg to fit in the hole tracks the fact that the peg is square and the hole round through a whole set of variations elsewhere in the circumstances. In no particular case will the squareness and the roundness as such figure in the full story of the multitude of interactions

---

[15] This sense of 'tracking' is weaker than Robert Nozick's in *Philosophical Explanations*, Oxford University Press, 1981, p. 178. It involves an analogue of his sufficiency condition but not of his necessity clause. On related matters, see Fred Dretske, 'Misrepresentation', in R. J. Bogdan (ed.), *Belief*, Oxford University Press, 1986.

[16] For earlier and less transparent accounts of the idea of distinguishing programme from process explanations, see the distinction between normalizing and regularizing explanations in Philip Pettit, 'Broad-minded Explanation and Psychology', in Pettit and McDowell, op. cit., and Frank Jackson and Robert Pargetter, 'Causal Statements', *Philosophical Topics*, forthcoming. We are indebted here to a number of discussions with Robert Pargetter.

which stop the peg from fitting into the hole, but the fact of squareness and roundness ensures, though not causally, that there is some very complex set of interactions which stops the peg from fitting into the hole.

Or take the case where we explain the shattering of a glass vessel in terms of the increase in temperature of the gas inside it. The increase of temperature is just the increase in the result of adding up the kinetic energy of all the molecules and dividing by the number of molecules, and clearly that increase does not itself cause the shattering. It is the impact of a relatively very small number of molecules on the walls of the vessel which actually breaks the glass. Nevertheless, we properly count citing the increase in temperature as explaining the shattering, for the increase programmes the shattering. There are many ways the increase in temperature might have been realized. All of them (again barring extremely implausible ones) would have caused the shattering, and one actually did cause it.[17]

Functional state explanations are a species of programme explanations. Take the simplest (and least informative) kind of functional state explanation. A glass's breaking is explained in terms of its being fragile. A familiar puzzle about this explanation has been that it is the categorical basis of the glass, not its fragility, that causes the breaking (along with the impact on the floor or whatever). Our answer to this puzzle is that when we explain the glass's breaking in terms of its being fragile, we are abstracting away from the particular basis of its fragility. We are saying that any basis for fragility would have combined with the impact to cause breaking. One basis did the work—the basis which was in fact the state which played the fragility role in the glass—but also a whole range of bases could have done the work of combining with the impact to cause breaking: it is like the explanation of the conductor's annoyance in terms of someone coughing. That explanation says more than that it was Fred's coughing, if it was indeed Fred who coughed, that caused the annoyance. It says that any of a whole range of members of the audience coughing would have caused annoyance in the conductor. If it turned out that it was crucial that it was Fred—he always coughs, and that was what really got the conductor's goat—then we would have to replace someone's coughing by *Fred's* coughing as the explanation of the conductor's annoyance.

Functional state explanations, then, are programme explanations. They give us a highly relational property, the property corresponding to the specification, which picks out a range of states satisfying the functional specification such that one state in the range did the causing, but the others would have had they been realized. The result is explained as a feature that tracks the satisfying of the functional specification. Our point here is a natural corollary of the familiar point about multiple realizability and functionalism. What matters for being in a certain mental state is not what

---

[17] We believe that some of the puzzles about teleological and evolutionary explanations can be illuminated by seeing them in terms of programme explanations, but that is another story.

realizes the functional role but the fact that the role is realized in one of the many ways that it might have been. In the same way, what matters for the causing of a certain result may be, in one sense, the functional role filled, not what fills it. Of course, in another sense, what matters is what causally produces the result; but in the sense at issue it is what does or would have produced the result, and functional state explanations, like programme explanations in general, tell us about the range of states that do or would produce the result without telling us which state in fact did the job. Just as the functionalist theory of mind tells us that what matters for being in mental state M is the fact that role F(M) is filled, not what fills it; so programme explanations tell us that what matters for some result R is the fact that some highly relational state of affairs is satisfied, not what satisfies it.

There can be more and less informative ways of picking out the range of states which would have produced a certain result, and so more and less informative programme explanations, and, in particular, more and less informative functional state explanations. Explaining breaking on being knocked in terms of fragility is a good example of a less informative functional state explanation. For in the explanation, the range of states that would have produced a certain result is specified as the states satisfying the functional role definitive of fragility, and that role is specified in terms of being knocked as input and breaking as output. But we already know that knocking led to breaking. All we are getting by way of additional information is that a range of categorical bases would have caused the breaking; it was not, that is, a peculiarity of the basis that actually did the causing which was crucial for the breaking. This is not no information, which is why we count dispositional explanations as explanations, but it is not a lot of information, which is why we count them as marginal explanations.[18]

However, the specification of functional role need not always be so close to what we already know. I observe that Fred's sighting of that cup causes him to reach for it, and explain his behaviour in terms of his belief that it contains water. This explanation says that a range of states playing the belief that-that-cup-contains-water role would have caused him to reach for it, and that one of them did cause him to. This is considerable additional information over and above what I know merely from the fact that he reaches for the cup and does so because of how things are with him. The functional role definitive of that content is highly complex, is consistent with his not reaching for the cup (as, for instance, in the case where he does not want water), and involves the causal antecedents of the internal state. Hence, although the glass which breaks because, in part, of how it is

---

[18] Our discussion of programme explanations is designed to be relatively neutral as between various views about causal explanation, but here is one place where the influence of David Lewis, 'Causal Explanation' in his *Philosophical Papers*, vol. II, Oxford University Press, 1986, shows.

constituted is by that very fact fragile, the agent who reaches for the cup because, in part, of how he is need not have the belief that that cup contains water.

We introduced programme explanations as our response to the concern that it is not having a certain content (broad or narrow) that enables a belief to drive behaviour: that it is instead the neurophysiological properties of beliefs which are causally efficacious, being the 'visible' properties in the chain of states mediating input and output. We now apply the distinction between programme and process explanations to the diagnosis of where the doppelgänger challenge goes wrong. Thus, though that challenge is directed to broad content only, if we are right, the one notion—that of a programme explanation—is the key to how the ascription of content explains behaviour, quite independently of whether the content is broad or narrow. There is nothing especially problematic about the explanatory value of broad content.

The mistake in the doppelgänger challenge

The challenge is to say what—in view of the similarity in behaviour of doppelgängers—an explanation by broad content could add to an explanation by correlative narrow content plus environmental detail. Our answer turns on the point that explanations by broad content are programme explanations.

We will introduce the crucial idea with an example. An object X is pressed into damp clay. The clay is allowed to harden with the impression of X's shape recorded in it. Object Y is then placed in the impression and fits perfectly. How should we explain this fact? One explanation might run somewhat as follows. X left a certain impression, say, a round one of five centimetres in diameter, in the clay. Y fitted snugly because it too is round and five centimetres in diameter. This is in terms of a narrow state of the clay, that of having a round, five centimetres impression in it, an impression our piece of clay shares with all its doppelgängers. Another explanation might be that the impression was caused by an object, as it happens X, which is the same shape as Y. Having an impression caused by an object which is the same shape as Y is a broad state of the clay, because it is not one it shares with all its doppelgängers, and hence this second explanation is one in broad terms.

It would clearly be a mistake to hold that the second (broad) explanation was redundant in the presence of the first (narrow) one. The second gives us explanatory information not contained in the first. It tells us that the particular shapes and sizes were not important, it was their all being the same that mattered. Had X and Y both been ten centimetres square instead of both being five centimetres round, we still would have had a snug fit. Of course, it is also true that the first contains information not in the second, for

the first tells us which particular shapes and sizes were causally involved. Neither explanation makes the other redundant; each has its own distinctive explanatory role. A doppelgänger challenge here would emphasize the fact that Y will fit just as snugly into all the duplicates of our piece of clay as it does in the original piece, despite the fact that the impressions in the duplicates do not all have the property of being caused by an object which is the same shape as Y. But this is simply irrelevant in the context of a programme explanation. The doppelgänger point—which is perfectly correct as far as it goes—tells us that if we keep constant the particular shape of both Y and the impression, we will get that same result of a snug fit, but clearly this in no way undermines the fact that if we vary the particular shape of Y and of the object causing the impression but keep constant their relationship, it is also true that we will get the same result of a snug fit. The crucial point is that the explanation in terms of the impression being caused by an object of the same shape as Y tells us the sense in which the particular shapes do not matter. The doppelgänger point draws our attention to the (different) sense in which the particular shapes do matter.

We say the same about explanation by the ascription of broad content. Although the environment will have caused some narrow state in me—let us suppose, for convenience of reference, that it is a belief with a narrow content expressible in English—which will in turn have caused a piece of behaviour by me, the particular narrow state does not matter, in the following sense. There will be an aspect of my behaviour, say, it's being a reaching for that cup, which remains constant under a whole range of possibilities concerning which narrow state does the causing, provided the cup acts on me in the right way for me to have a belief described as being about the cup. The cup will be a certain colour and shape, a certain distance from me, in a certain direction, illuminated in a certain way, and so on and so forth. All in all, the cup will present itself as, say, the F. These facts together will cause me to believe that the F is desirable under some aspect, and so to reach for the F, and, as the F is that cup, to reach for that cup. But equally had the cup been a different colour and shape, a different distance and direction from me and so on, in short, the G and not the F, I still would have reached for it. For I would in that case have believed that the G is desirable and, consequently, have reached for the G; but in that case the G would have been that cup, and so it still would have been true that I reached for that cup.

The distinctive explanatory role of broad content can be brought out in terms of the multiple realizability of broad content by narrow content and environment. For a given broad content B there will be a number of ways of realizing that content by the appropriate combination of narrow content N and environment E, say: N_1 & E_1, N_2 & E_2, ... One of these ways, say Na & Ea, will be the actual way B is realized. Now each of the N_i & E_i will explain and predict different behaviour in the subject, but it may be that

there is a common thread T running through these different pieces of behaviour. In this case ascribing B explains and predicts T just as well as ascribing Na & Ea, and does something distinctive besides—it tells us that it did not matter as far as getting T goes that it was Na & Ea that was actual instead of, say, N_2 & E_2. It might be objected that to say that Na & Ea in fact realizes B is to say that Na & Ea constitutes (as things in fact are) B. So how can ascribing Na & Ea not explain all, and possibly more besides, that ascribing B does? Our answer should by now be clear. Fred's coughing may, as things in fact are, constitute someone's coughing, and yet explaining the conductor's annoyance in terms of the fact that someone coughed is not subsumed in an explanation in terms of the fact that Fred coughed. Similarly, Na & Ea may as it happens constitute B, yet an explanation in terms of B is not subsumed in an explanation in terms of Na & Ea.

In sum, programme explanations are explanations that give a range of states which would have produced the result we want explained. Explanation in terms of broad content picks out the range of states via *inter alia* the role of the environment in causing the states. The causally efficacious properties of the states will be neurophysiological properties, and different environmental stimuli will produce neurophysiologically different states, and so different behaviours. However, it may be that if we keep fixed some feature of the environment, then the different behaviours display a common feature despite variations elsewhere. In such a case, the feature of the environment may be part of a programme explanation of the behavioural constancy. Our claim is that explanations in terms of broad content employ a relationship to the environment—thus, the context-bound nature of broad content—as part of what singles out the range of narrow states that would produce, and one of which does produce, behaviour manifesting the feature the broad content explains. We said earlier that programme explanations concern what remains constant under variation. The explanatory point of ascribing broad content is that keeping a relationship of agents to external factors constant through variations inside those agents may preserve as a constant some feature of their otherwise varying behaviour. Some feature of their behaviour may track the relationship to external factors.[19]

CONCLUSION

We started from the fact of broad content, from the fact that some intentional states have the content they do have in virtue of how things are around their subjects in combination with how things are inside their subjects. Our question was how functionalists should respond to this fact. Our answer was that they should respond in the least disruptive fashion possible, namely, by sometimes including outside factors in the inputs and

---

[19] *Ceteribus paribus* and within limits as always, but particularly where psychological explanation is concerned.

outputs that figure in their story. We defended this answer against two widely entertained objections. The first is that the fact of broad content undermines the whole picture of the mind as something internal (a picture which brings functionalist and Cartesian into unholy alliance). Our reply was that functionalists at least can be neutral on whether the mind is internal in the relevant sense. For, by choosing what we called the role state option, they can make an intentional state's content an essential property of it, and so make a state with broad content to be such that its very identity, not just how it may be described, is linked to the surroundings of the subject which has the state.

   The second objection is that broad content does not play the kind of explanatory role needed to justify incorporating it into the functionalist story in our simple (minded?) way. It is, as the doppelgänger challenge shows, narrow content which plays that kind of role, and accordingly we should be dualists about content. Narrow content does the explaining of behaviour in a way right for inclusion in the (narrow) functionalist story. Our reply to this objection was to argue that properties may be causally explanatory properties without being causally productive or efficacious ones. These properties programme the result to be explained, rather than actually bringing it about, and are the properties appealed to in what we called programme explanations. We argued that programme explanations are very common, being the kind of explanations offered when we explain in terms of highly relational properties, and that explanations in terms of content, be it narrow or broad content, are programme explanations. This fact enabled us to identify a distinctive explanatory role for broad content, a role not filled by the correlative narrow content, and so to justify broad content's inclusion in its own right in the functionalist story about content.[20]

Research School of Social Sciences   FRANK JACKSON
Australian National University   PHILIP PETTIT
Canberra, ACT 2601
Australia