Subject, Thought, and Context

EDITED BY

PHILIP PETTIT

AND

JOHN McDOWELL

CLARENDON PRESS · OXFORD

1986

CHAPTER 1

BROAD-MINDED EXPLANATION AND PSYCHOLOGY*

PHILIP PETTIT

1. Introduction: The Issue of Explanatory Dispensability

Broad-minded explanation is explanation which calls on beliefs, desires, and the like—in general, on intentional states—that are themselves broad or wide. The question I wish to raise in this paper is whether such explanation is ever psychologically indispensable; in particular, whether it is ever more than a stand-in for an account which only invokes narrow psychological states.

The broad–narrow divide is drawn at different places, depending on where the boundary is thought to fall between a person's 'core', as some at least will take it, and his context or surrounds. By some accounts the boundary falls at the surface of the body; by others it falls somewhere nearer the brain; by still others, it cannot be located in such physicalistic terms. We can leave open the issue between these accounts, however, defining the broad–narrow distinction as follows. A token or particular state is narrow if the presence of that type of state is guaranteed by the context-independent character of the subject; otherwise the state is broad.

This definition is meant to be neutral on the issue dividing dualists and materialists. For all that it says, the context-free character of the subject, or indeed the character of any part of the world, may be physical or non-physical. The idea is that, whatever the stuff out of which people are composed, their narrow states are those which supervene on how it is with the subjects, independently of the nature of their environment, while their broad states are those which

 I am indebted to a number of people for comments on earlier versions of this paper. In particular I must mention Jeremy Butterfield, Michael Devitt, Peter Forrest, Peter Godfrey-Smith, Frank Jackson, Mark Johnson, Graham Macdonald, John McDowell, Peter Menzies, David Papineau, Huw Price, Christie Slade, Jack Smart, Kim Sterelny, and Stephen Stich.

supervene on how it is both with the subjects and with their environment. The narrow states cannot cease to exist without a change in the appropriate context-free base; keep the base constant and they are bound to survive. By contrast, the broad states may cease to exist without any change in that inner space. A change of context, just on its own, can cause one of the states to disappear.

Under this account of the broad–narrow divide, any narrow type of state that I exemplify will be realized also in a *Doppelgänger*: that is, in someone who, though inhabiting a distinct and perhaps very different environment, is indistinguishable from me in context-independent character.[1] This is not true of every broad state: depending on the variation in context, it may or may not be replicated in my unearthly twin. We have on offer here a thought test for the 'spread' of a state. Given the state, consider whether it is replicated in all the subject's possible *Doppelgänger*. If it is, the state is narrow; if it is not, the state is broad.[2]

I am going to assume that intentional states at least include some broad states, and that broad intentional states figure in psychological explanation. The traditional paradigm of such a state is knowledge. If someone knows that *p*, when the truth of that proposition turns on things outside himself, then he is not in a narrow state. Where it was not the case that *p*, no one would know that *p*, for what is known must actually obtain.

Knowledge is broad because it requires context to be such that its content is true. Other intentional states such as belief and desire are less demanding in regard to truth, but may still involve requirements on the world which ensure that they are broad. Consider the belief that *p*, where the content or meaning expressed by the '*p*'-sentence is determined in part by how the world is: this may be because the sentence contains a proper name, a natural-kind term, a demonstrative,

[1] Twin Earth, under Hilary Putnam's original presentation of this sort of idea. See his 'The Meaning of "Meaning"', in his *Mind, Language and Reality* (CUP, Cambridge, 1975).

[2] This means, I accept, that for someone to believe that he is such and such is for him to token a broad state; his twin will token the different type of state that he, the twin, is such and such. See Harold Noonan, 'Russellian Thoughts and Methodological Solipsism', in Jeremy Butterfield (ed.), *Language, Mind and Logic* (CUP, Cambridge, forthcoming). I am consoled by John McDowell's observation that nevertheless the twins will each instantiate the state of believing oneself to be such and such.

or whatever. This belief will be broad because it will require the world to be such that its content remains the same, determining in particular the same truth condition: the same function from possible worlds to truth values.[3]

There are sources of breadth other than the facts just mentioned: viz. that the world determines the truth values of intentional contents, and that it sometimes determines even their truth conditions. One is the fact, alleged under some interpretations of Wittgenstein, that when an intentional state involves a rule-following disposition towards its content, then it requires that the community of the bearer be and remain of a certain character.[4] Another is the fact that if we assign a belief or whatever on the basis of what someone says, and if we overlook idiosyncrasies of understanding, construing his words in their standard community sense, then which belief we ascribe will depend in part on the practice of that community, in particular on how it uses those words.[5]

Back then to our question. Assuming that there are some broad intentional states, and that they figure in psychological explanation, are they ever indispensable in explanation? The question is crucial, since such indispensability is likely to be made a hallmark of genuinely mental intentional states. Broad intentional states may be of social or semantic importance, but, failing explanatory indispensability, they will not be counted in any sober inventory of the psychological realm.

They will be treated, in all likelihood, in the manner to which knowledge has become accustomed. The received analysis of knowing that *p* factors it into two components: on the one hand, believing that *p*; on the other, its being the case that *p* and that the belief that *p* is justified or whatever. This decomposition presents knowledge as a hybrid state, involving the psychological component of belief and a non-psychological environmental correlate. The belief component

[3] See Gareth Evans, *The Varieties of Reference* (Clarendon Press, Oxford, 1982). For a survey of the theories of reference involved see Nathan U. Salmon, *Reference and Essence* (Blackwell, Oxford, 1982).

[4] See my paper 'Wittgenstein, Individualism and the Mental', in Paul Weingartner (ed.), *Epistemology and Philosophy of Science*, Proceedings of the Seventh International Wittgenstein Symposium (Hölder–Pichler–Tempsky, Vienna, 1983).

[5] See Tyler Burge, 'Individualism and the Mental', *Midwest Studies in Philosophy*, Vol. IV (1979).

will do all the explanatory psychological work, on this account, and so knowledge will be denied a place among the furniture of the mind.

Intentional states figure in at least two sorts of psychological explanation: the explanation of action and the explanation of the appearance of other intentional states. In this paper I shall focus primarily on the explanation of action. I shall be arguing that while broad-minded explanation cannot be indispensable under the standard account of action-explanation, it is indispensable under an alternative and superior account. My claim is that if action-explanation invokes broad intentional states—as I am assuming it does—then it calls upon them indispensably. The force of the explanation is bound to be lost, if reference to those states is eliminated.

It would be possible for someone to concede my claim and then go on to say that this does not in itself establish the credentials of broad intentional states. What it shows, he will say, is that if our ordinary pattern of action-explanation is to be taken as a pointer to the nature of the mind, then broad intentional states have to be countenanced. But it also has to be demonstrated, he will add, that this pattern should be taken seriously: that it is not based, for example, on a misconception of the sorts of things that require psychological explanation.[6] I will not seek to provide such a demonstration in this paper. My argument is addressed to those who assume that the ordinary explanation of action is a fitting vantage point from which to plot the contours of the mind.

The argument is conducted as follows. In section 2 I set out three points of agreement between the standard account of action-explanation and the account which I shall be offering. In section 3 I characterize the standard account and in section 4 I show why it would force us to see broad-minded explanation as dispensable. In section 5 I present and provide support for my alternative to that account and in section 6 I explain why this, in contrast, gives no ground for thinking that we can dispense with broad explanatory states. Finally, in section 7 I give a brief characterization of the view of mental states implicit in my approach; in particular, I try to show where it differs from the orthodox functionalist one.

[6] Paul Churchland and Stephen Stich both take the view that ordinary action-explanation is misconceived. See Churchland, 'Eliminative Materialism and Propositional Attitudes', *Journal of Philosophy*, lxxvii (1971), and Stich, *From Folk Psychology to Cognitive Science* (MIT Press, Cambridge, Massachusetts, 1983).

2. *Some Points of Agreement about Action-explanation*

There are three important points of agreement between the standard account of action-explanation and the account which I shall be offering. These I need not defend in detail, precisely because they are agreed, if not on all sides, at least on both of the sides which we shall be examining.

The first is that action-explanation is reason-giving. The intentional states which it explicitly or implicitly ascribes to the agent give a reason for the action under its description as explanandum. Suppose the action is described as *A*-ing. The states invoked in the explanation constitute a reason for the person's *A*-ing if and only if they include a pro-attitude towards a certain sort of situation plus a belief that by *A*-ing the agent can bring about that situation.[7] I put on the car heater because (i) I wish to have a clear windscreen and (ii) I believe that putting on the heater will realize that goal. The pattern is familiar and uncontentious.

The second point of agreement between the different accounts is that action-explanation is cause-giving as well as reason-giving: specifically, that the complex of intentional states which gives a reason for the action also causes it to occur. If the agent's being in those states explains his action, then it is not enough that the states constitute a reason: despite their service in this regard, the action might be caused by something else. In order to have explanatory force, the complex of states must produce the behaviour as well.[8]

The third and final point of convergence is that action-explanation is pattern-giving, a pattern being a non-accidental regularity. This point is not often spelled out, but it is implicit in most accounts and is in any case persuasive.[9]

The idea is not that the explanation of action is guided by antecedently formulated principles, enunciating general regularities. It is only that the person who explains an action *A* by the presence of an

[7] This account is modelled on Donald Davidson's classic 1963 paper 'Actions, Reasons, and Causes', reprinted in his *Essays on Actions and Events* (Clarendon Press, Oxford, 1980). See in particular pp. 5 and 8–9.

[8] See Davidson, ibid., pp. 8–19. I ignore the further requirement that the intentional complex must produce the action in a non-deviant way.

[9] See Dagfinn Føllesdal, 'The Explanation of Action', in R. Hilpinen (ed.), *Rationality in Science* (Reidel, Dordrecht, 1980), 237: 'There must be some regularity, which we can at least roughly specify, that connects the reason with the action.'

intentional complex *I* must be able to argue for a pattern relating *I*-type profiles to *A*-type actions. The idea is meant to apply to any event-explanations. To explain an event *E* by a cause *C* one must be able to make out a general connection between *C*-type and *E*-type events. Or so the story goes.

Strictly, there are three strands to the story. The first is that for any causally related pair of events *C* and *E*, some relation such as the following must obtain: that the first necessitates the second, that it makes it more probable than not, or at least, excepting determination, that it makes it more probable than it would have been in an otherwise similar situation where the first had not occurred. Unless some such relation obtains between causes and effects, it is unclear why we concern ourselves so much with causes; in particular, it is unclear why we think of bringing about effects by bringing about their causes.

The second strand of the story is that a relation of the kind envisaged does not float free of the other general properties of the events and their situation. For whatever reason, we rule out the possibility that it should hold in only one of two otherwise similar set-ups. The relation is supervenient on other general properties. Keep those the same and it too must remain constant.

What is entailed by these two strands is that if someone causally explains *E* by *C*, then he commits himself to the truth of a universal principle. This principle is, at its weakest, that in any exactly similar situation the counterpart of *C* necessitates or in some sense probabilifies a counterpart of *E*. If this were all that held, however, we could not really say that causal explanation was pattern-giving. A pattern is given to us by an explanation only if it is picked out under an informative characterization and only if, so characterized, it is independently plausible. The pattern adumbrated here satisfies neither of these conditions.

The third strand of the story makes up the gap. It is agreed on all sides that when we explain an action, or indeed any event, we are in a position to offer an informatively characterized and independently plausible principle in support of the explanation. We can conceive of that principle as enunciating, or at least as providing evidence for, the general pattern entailed; in either case we can say that the pattern is given. The reason for saying that the principle may only provide evidence for the pattern is that the concepts used in the explanation,

and therefore in the principle, may not be suitable for the precise formulation of the pattern.[10]

Notice that in envisaging the possibility that a principle may enunciate or just provide evidence for a pattern, we are conceiving of patterns in a distinctive sense. In this sense a pattern is something ontological, not a linguistic or epistemic entity like a principle or law or generalization. As a result it is a perfectly determinate, though not necessarily a deterministic sort of regularity; it is not subject to any vagueness or to any open-ended *ceteris paribus* clauses. Such imperfections may attend our efforts to express patterns but not the patterns themselves.[11]

The last strand of the pattern-giving story is borne out in our explanatory practice. For any event-explanation we offer, we can readily formulate a plausible principle linking the cause-type event to one of the effect-type. The stone broke the window, we say, and we go on easily to generalize about the effects on glass of stone-like objects thrown with a certain force.

If an explanation is to have its proper force, then cause and effect must be described so that we can generalize in this manner. Suppose we are told that the event which Johnny witnessed caused the branch to fall from the tree—or, even worse, that it caused the event which upset Aunt Mary. This does not explain, except in a purely empty way. It may give us grounds for believing that the event is intelligible, but, not suggesting a generalization, it does not serve the distinctive explanatory task of enabling us to find the event intelligible. The difference between finding an event intelligible and merely believing it so is immense. It is like the difference between finding a

[10] See Davidson, op. cit., p. 16, on the principle which most of us will offer in explaining the breakage of a window by a rock. 'A generalization like, "Windows are fragile, and fragile things tend to break when struck hard enough, other conditions being right" is not a predictive law in the rough—the predictive law, if we had it, would be quantitative and would use very different concepts. The generalization, like our generalizations about behaviour, serves a different function: it provides evidence for the existence of a causal law covering the case at hand.'

[11] Might a pattern resist the sort of formulation which we would describe as law-like enunciation, and yet answer to a principle which provides evidence for it? Those who hold that causal explanations do not entail law-like connections would presumably say that it might. They could thereby endorse the points made in this section. For an example of someone who holds that causal explanation does not entail a law, see J. R. Searle, *Intentionality* (CUP, Cambridge, 1983), 120–1.

joke amusing and inferring that it is amusing from the reactions of others.

I conclude that event-explanation in general, and action-explanation in particular, must give a pattern as well as a reason and a cause. But what sort of pattern is provided in the action case? What type of principle is adumbrated in the explanatory invocation of intentional states? The way to get at the principle on offer in any explanation is to ask how we would generalize it to other cases and how we would modify those generalizations in response to various objections. As it happens, Paul Churchland has already gone through this exercise for the case of action-explanation.

His analysis goes as follows, where '*X*' ranges over agents, '*O*' over states of affairs, and '*A*' over actions: for all *X*, all *O*, and all *A*, 'If (1) *X* wants *O*, and (2) *X* believes that *A*-ing is a way for him to bring about *O* under the circumstances, and (3) there is no action believed by *X* to be a way for him to bring about *O*, under the circumstances, which *X* judges to be as preferable to him as, or more preferable to him than, *A*-ing, and (4) *X* has no other want (or set of them) which, under the circumstances, overrides his want *O*, and (5) *X* knows how to *A*, and (6) *X* is able to *A*, then (7) *X A*-s.'[12]

I propose to accept this analysis, with one amendment. I would prefer to replace at least clause (4) with a simple *ceteris paribus* condition. The reason is that spelling out the clause as Churchland does suggests that we can independently determine that other things are equal. This is misleading. We can only tell *ex post* that a want is overridden. And in any case this is not the only story available to explain an agent's failure to do *A* under the circumstances described by the other clauses. We might want to say that the want was displaced by some other desire, where this does not have the same deliberative implications as saying that it was overridden. Or we might want to postulate the sort of breakdown often associated with weakness of will.

It is not a scandal that the principles on offer in action-explanations should be as open-textured as I am suggesting. The principles need not be taken to enunciate determinate patterns in nature, only to provide evidence of their existence. We should not be surprised to find that such principles cannot be explicated in detail; and this, no matter what view we take of action-explanation.

[12] Paul Churchland, 'The Logical Character of Action-explanations', *Philosophical Review*, lxxix (1970), 221–2.

3. *The Standard Account of Action-explanation*

For all that has been said so far, we still have no story about why it is that mention of a cause, even if it provides evidence of a pattern, can have explanatory force: that is, can enable us to find the effect intelligible. We may agree that such mention of a cause is necessary for event-explanation. But why is it sufficient? The standard account provides the orthodox answer.

The story is that to find an event intelligible is to see it as a particular instance of how the world works generally. It is to recognize the event as unexceptional, as the routine operation of a recurrent process. Thus, so the account goes, to explain an event is to provide the materials for such accommodation to the singular. It is to regularize the event in question.[13]

Fully articulated, the predicate 'explains' has more than the two places required for the explanans *p* and the explanandum *q*. In order to specify an appropriate generalization, it must leave room for mention of the possibilities by contrast with which *p* is explanatory and *q* is explained: say, *o* and *r* respectively. And, more important for our present purposes, it also has to allow space for the assumptions against the background of which intelligibility is revealed: say, *s*. It is shorthand to say simply that *p* explains *q*. Properly speaking, we should say that *p* (rather than *o*) explains *q* (rather than *r*), given *s*.[14]

The standard account of explanation tells us something about the assumptions, *s*. On this account, what is assumed by way of background to event-explanation is that the world works to a causal order, embodying mechanical, if sometimes only probabilistic, patterns. The idea is that, given such a background, an event is rendered intelligible when we are shown that it is part of the dispensation assumed.

The standard account is not implausible, and it certainly holds of

[13] The view is sometimes known as the subsumption theory. See Christopher Peacocke, *Holistic Explanation* (Clarendon Press, Oxford, 1979), 154. I do not like that description, since it suggests that only in this form of explanation does a general principle play a role.

[14] See Bas C. Van Fraassen, *The Scientific Image* (Clarendon Press, Oxford, 1980), Chapter 5.

some event-explanations. It makes sense of two points admitted in the last section: that the explanation of an event mentions the cause, and that in the mode of mention it points us towards an independently plausible and informatively characterized pattern. This is precisely what we should expect if the point of the explanation is to display the event as an example of the mechanical order in things.

We saw in the last section that the principle deployed in an explanation may enunciate or, because of its conceptual limitations, may merely provide evidence of a pattern. This means that there will be two ways of regularizing an event: one, the proper way, by enunciating the pattern under which the event is subsumed; the other, the proleptic or anticipatory way, by merely providing evidence of its existence. Although it is less significant than proper explanation, notice that the proleptic sort is much more satisfactory than the empty account in which no principle is put on offer, as when we say that the event Johnny witnessed caused the branch to fall.

How does the standard account of explanation fare with action-explanation in particular? On the face of it, perfectly well. Folk psychology provides us with principles conforming to the belief–desire schema given in the last section. These can be taken as at least evidencing certain patterns in nature, thereby allowing proleptic regularization. When one of them is applied, as in the explanation of an action, then the instantiation of the antecedent—an intentional profile—is agreed to be the cause of the action explained. Thus we can coherently cast an action-explanation as an attempt to accommodate the action within a mechanical picture of the world, an attempt to regularize the action by displaying it as an instance of how the world works generally.[15]

[15] See for example J. A. Fodor, *Psychological Explanation* (Random House, New York, 1968), 45: 'our current account of causal explanations requires only that we demonstrate that a certain action occurs whenever specified conditions are satisfied'. Even Robert Cummins, who insists on the distinctive nature of psychological explanation, endorses such an account of action-explanation, and sees it as standard. See his *The Nature of Psychological Explanation* (MIT Press, Cambridge, Massachusetts, 1983), 14. In his view psychological explanation is distinctive so far as it follows a certain strategy for making sense, not of actions, but of capacities. The same is true of John Haugeland, 'The Nature and Plausibility of Cognitivism', *Brain and Behavioral Sciences*, ii (1978), 215–60. Some authors who endorse the regularizing account are cautious enough to insist that action-explanation involves regularization plus something else. See Colin McGinn, 'The Structure of Content', in Andrew Woodfield (ed.), *Thought and Object* (Clarendon Press, Oxford, 1982), 255, n. 3.

A striking feature of action-explanation is that the covering principles—those conforming to the belief–desire schema—are knowable a priori. One does not have to search around for inductive evidence to learn that if it is desirable that p, and if it happens that not p, and if one can ensure that p by A-ing, then other things being equal one ought to A. Equally, moving to the third person, one does not have to rely on inductive premisses to establish that if someone desires that p, believes that not p, and believes that by A-ing he can ensure that p, then *ceteris paribus* he As. Such a principle gives expression to our conception of what it is to believe and desire things. Understand that conception and you will be in a position to see that the principle is true.

Is its a priori character a difficulty for the claim that the principle serves to regularize action? It is not a metaphysical difficulty, for sure. There is no reason why the principle on the basis of which an event is regularized and explained—even properly, let alone proleptically—should not be a priori. That it is a priori may mean just that it is so deeply embedded in our web of belief—and our way of speaking—that we cannot, at least as we are at present, envisage circumstances under which we would give it up. Why should principles that enjoy such an epistemic status be prohibited from serving in regularizing explanations?

Still, the a priori character of the principles may be thought to raise a methodological worry. With such principles, the failure of the consequent to follow cannot be allowed to disprove the law: it must be taken to show that the antecedent is not after all satisfied. This being so, it appears that the antecedent is not verifiable without verifying the consequent. And that would certainly be a problem; it would introduce a destructive circularity.

But the problem is overstated. It does not follow, from the fact that the falsity of the consequent would undermine the ascription of the antecedent, that the antecedent is not independently verifiable. We may have lots of independent evidence for ascribing the antecedent state: evidence related to the circumstances of the agent, his background training and skills, and his other actions and utterances.

If this gets rid of our first methodological worry, however, it leaves room for a second. The worry is that although the antecedent may be independently verifiable in a particular instance, we would not ascribe it in any instance unless we generally found that the

consequent followed.[16] Consider a parallel. We may be able to judge in an individual instance that an object is fragile and a shock it receives severe, without being able to see whether the shock shatters the object. The antecedent of the following a priori principle is therefore independently verifiable: any fragile object which is submitted to a severe shock breaks. But we may still feel qualms about that principle, especially as a principle giving us an explanatory regularity. There is a conceptual connection between fragility and breaking which makes the principle suspiciously convenient.

The difficulty can be stated persuasively in the fragility case. The a priori covering principle, independently verifiable though its antecedent may be, does not have the same regularizing force as a principle which invokes the molecular structure of the object instead of its fragility. Molecular structure can be determined without reference to whether the object breaks under severe shocks, and there is nothing suspicious about the regularity which the principle invoking it reveals in the breakage. There is something shady in comparison about the regularization effected by the fragility principle.

But if the difficulty is clear in the fragility case, so is the resolution. The explanation by means of the a priori principle can be held to go proxy for an account in terms of molecular structure. Referring to fragility is referring to whatever intrinsic property it is—say, such and such a molecular structure—that causes things to shatter under certain pressures. Explaining by reference to fragility is not giving the ideal regularizing account, but only doing the best possible under ordinary conditions of ignorance: namely, indicating the form which the ideal account should take. Where the molecular structure account is a proper explanation, the fragility account is a proleptic one.

The person who defends the standard account of action-explanation can respond in similar fashion. The intentional profile invoked in action-explanation is subject to more result-independent conditions of ascription than the property of fragility. But if it is thought to be still suspiciously convenient, then the defender of the standard account can say that its explanatory force derives from the fact that it goes proxy for some intrinsic property of the agent. The picture is that the ideal principle to explain an action would invoke something like a neurophysiological state or complex of states, and

[16] This sort of difficulty is presented in Peacocke, *Holistic Explanation*, 147–55, as the failure to meet a certain 'conjunction restriction' on explanations.

that the intentional profile actually called upon stands to this as fragility relates to molecular structure.[17]

This supports our earlier suggestion that on the standard account of action-explanation, the principles will be taken to play a proleptically explanatory role, not a proper one. They will be seen as principles which merely provide evidence of the patterns in virtue of which the explanation works, not as principles that strictly enunciate those patterns: this, at least, if our concepts of pattern and principle are the ones deployed.[18]

Given this picture, it is natural to ask about the relation between belief–desire principles and the more fundamental laws which do enunciate the relevant patterns: this, assuming that the patterns can be formulated in law-like terms.[19] The principles obtain on the basis of those laws. But are they then reducible to them? Or do they obtain in virtue of those other laws, but in such a way that the laws cannot be used to define necessary and sufficient conditions for their truth? Are they merely supervenient, in other words, on those laws?

The supervenience story is the most common one in the recent literature, and it is usually elaborated as follows. Suppose that an agent instantiates a belief–desire principle, manifesting the intentional profile described in the antecedent and going on to display the behaviour mentioned in the consequent. The tokens of the two types of state in question—mental and behavioural—may also be capable of being typed differently—say, neurophysiologically; and if they are not susceptible to such typing themselves, they may be associated with a pair of token states that are. What happens on the supervenience picture is that in every instance where a belief–desire principle applies, it does so because of the satisfaction of another sort of law—say, a neurophysiological one—by the appropriate couple of token states; and the law on the basis of which it applies may vary more or less wildly from instance to instance.[20]

[17] See for example Fodor, *Psychological Explanation*, 34–6. For a comment, see D. M. Armstrong, 'Recent Work on the Relation of Mind and Brain', in *Contemporary Philosophy: A New Survey*, Vol. 4 (M. Nijhoff, The Hague, 1983), 56–7.

[18] Obviously one might take every principle to enunciate some pattern, allowing patterns to be more or less determinate. In that case, one would formulate the point of this paragraph differently.

[19] The assumption may be rejected. See n. 11.

[20] The supervenience picture is common to a variety of philosophies. It is developed in one way by Davidson's 'anomalous monism'—see 'Mental Events', reprinted in *Essays on Actions and Events*; and in another by certain functionalist theories—see Colin McGinn, *The Character of Mind* (OUP, Oxford, 1982), 33–6.

Enough has been said to show how the standard account of explanation can be fitted to action-explanation in particular. There remains one question. We saw in the last section that action-explanation is not only cause-giving and pattern-giving, but also reason-giving. What relevance can attach in the standard account to the fact that the intentional profile invoked in explanation of an action must be a reason as well as a cause? If the explanation of an action makes it intelligible just by displaying it as a regular event—an instance of how the world generally works—then the fact that the cause invoked is also a reason must be strictly irrelevant. It is the regularization of the action, and that alone, which illuminates the event. But something more must be said. The standard account will be seriously deficient unless it offers some story as to why action-explanation trades in reasons.

The defender of the standard account will probably run the following line. In explaining action our working procedure is to assume that people are more or less rational: this, both in the way they form their beliefs and desires—something we have not considered so far—and in the manner in which they act, given those intentional states. The assumption of rationality seems to be heuristically indispensable, offering us our only clues as to the causes of behaviour. But the fact of relying on such an assumption means that the causes we unearth in any instance necessarily appear also as reasons for the action explained. It is because they are reasons—and because they are reasonable attitudes to ascribe—that we identify them as causes, though it is because they are causes that they explain.[21]

This final twist puts the standard account in nice perspective. The explanation of action uses the assumption of rationality in the enterprise of representing behaviour as part of the causally ordered world. That it does so means that the causes it picks out are reasons,

[21] Some defenders of the standard account reject the assumption of rationality, even as a heuristic principle, and they will have to tell a different story. See for example Stephen Stich, 'On the Ascription of Content', in Woodfield (ed.), *Thought and Object*, 153–206. Of those who endorse it, only some make clear that the assumption is strictly heuristic: for example, Brian Loar, *Mind and Meaning* (CUP, Cambridge, 1981), 130–2. That action-explanation explains so far as it gives us a cause is explicit in Fodor, *Psychological Explanation*, 41: 'in the sense of causal explanation here at issue an event has been explained if we can show that sufficient conditions for its occurrence have been satisfied'. See too David Papineau, 'Representation and Explanation', forthcoming in *Philosophy of Science*.

and that the covering principles it mobilizes predict the effects which such reasons dictate. Still, the fact that the causes and principles have this rational aspect is not essential to the explanations in which they serve. Those explanations explain by displaying regularity; it is incidental that the regularity displayed is rational.

4. The Standard Account makes Broad-minded Explanation Dispensable

If the standard regularizing conception of action-explanation is adopted, then it follows, given some plausible supplementary premisses, that the broad-minded explanation of action is psychologically dispensable. I want to show why this is so, before exploring my alternative account of action-explanation.[22]

Suppose that we are presented with an explanation of an action in which the explanans involves a broad intentional state. That state requires something of the agent's milieu, and if the requirement were not fulfilled, then the state would not be available to provide an explanation; this, moreover, even if there were no corresponding change in the context-independent character of the agent. We can envisage the state as a piece of knowledge or a demonstrative thought, for example. The question is whether the explanation it supports is necessarily dispensable in favour of an explanation—specifically, a psychological explanation—which involves only narrow states.

There are strictly two cases to consider: one where the explanandum is narrow, the other where it is broad. I shall speak only to the

[22] The section which follows provides foundations for what normally passes by assumption or, at best, is briefly stated: see for example Jaegwon Kim, 'Psychophysical Supervenience', *Philosophical Studies*, xli (1982), 65. See also Michael Devitt, *Realism and Truth* (Princeton University Press, Princeton, 1984), Chapter 6, and Stephen Stich, *From Folk Psychology to Cognitive Science*, Chapter 8. A common form in which the assumption appears is that if the existence conditions for the cause mentioned in a cause-giving explanation include some which at another level of analysis can be seen to be causally inert, then the antecedent has been misidentified. This is often quoted as a reason for shrinking the psychological antecedents of action to narrow dimensions. Even Daniel Dennett, whose views on explanation are in many ways congenial to those expressed here, is moved by this. See 'Beyond Belief', in Woodfield (ed.), *Thought and Object*, 13 and 26. See also Colin McGinn, 'The Structure of Content', same volume, p. 208, and David Lewis, 'Belief *De Dicto* and *De Se*', *Philosophical Review*, lxxxviii (1979), 526.

second, since if broad-minded explanation is dispensable in this instance, it will certainly be dispensable in the first also. The broad intentional state in the case imagined I shall ascribe by the sentence '*Ia*', where *a* is the subject and '*I*' designates the property attributed. The broadly described action which serves as explanandum I ascribe by the corresponding sentence '*Aa*'.

I have to demonstrate that in any case like this, where the fact that *Ia* serves to regularize the fact that *Aa*, the explanation can be provided by mention only of narrow psychological facts. The argument which I have to offer is usefully set out in three stages.

Stage 1

Because it is broad, we can legitimately decompose the fact that *Ia* as follows: the agent is such that, world willing, *Ia*. We can decompose the fact that *Aa* in a similar mode: the agent acts so that, world willing, *Aa*. I assume that one and the same condition is involved on the world's side. Let us ascribe this in the sentence '*Cw*', where *w* is the world and '*C*' designates the property involved in the condition. On the one hand, then, we have: the agent is such that if *Cw*, then *Ia*—for short, *I*a*; on the other, the agent acts so that if *Cw*, then *Aa*—for short, *A*a*.

It will be useful to illustrate the sort of decomposition envisaged. If '*Ia*' stands for '*a* knows that *p*', then '*I*a*' is '*a* is such that, world willing, he knows that *p*'. And what is it for the world to be willing here? At the least, the world must be such that *p*. Whether other conditions must be fulfilled, and if so what, is matter for debate.

Another example. Suppose that '*Ia*' stands for '*a* wants this cup'. '*I*a*' will then be '*a* is such that, world willing, he wants this cup'. Here the world will be willing just in case it contains the very cup in question. '*I*a*' then comes to: '*a* is such that, given a world where this cup exists, he wants this cup'.

The illustration of '*Aa*' and '*A*a*' follows similar lines. Suppose that '*Aa*' is '*a* kicks a goal'. '*A*a*' will then be '*a* evinces such narrowly characterizable behaviour that, given a world where the consequences are as here, he kicks a goal'. Again, suppose that '*Aa*' is '*a* grasps this cup'. '*A*a*' will be '*a* evinces such narrowly characterizable behaviour that, given a world where this cup exists, *a* grasps this cup'.

In virtue of our decomposition, it may appear that we can always find a narrow-minded regularization to replace a broad-minded one:

this, assuming that the *I**-state can be regarded as a psychological state. Where we might have invoked the fact that *Ia*, we can apparently call upon the fact that *I*a* and *Cw*. Since the *I*-state is broad, requiring that *Cw*, we know that necessarily it is the case that *Ia* if and only if *I*a* and *Cw*.

But we should be cautious, for any number of parallel decompositions are also on offer. For example, we might have decomposed the *I*-state into an *I*(@)-state, where *I*(@)*a* if and only if *a* is such that, if $1+1=2$, then *Ia*. Since $1+1=2$ in every possible world, we know that necessarily it is the case that *Ia* if and only if *I*(@)*a* and $1+1=2$. If we are prepared to say that anything which the fact that *Ia* regularizes can be regularized by the fact that *I*a* and *Cw*, we may be forced to add that it can also be regularized by the fact that *I*(@)*a* and $1+1=2$.

Where we have decompositional alternatives of this sort, we have to decide which provides the best regularizing explanation. If we decide in favour of one of them, then we know that the type of state involved there is primary and that the states involved in the alternatives are gerrymandered out of it. The primary type of state is the true explanatory kind, as we might say; the other types are forgeries.

How ought we to decide between decompositional alternatives? If they are alternatives for a proper form of regularization, then we will select the one whose principle enunciates the pattern in nature in the theoretically most satisfactory way: the one whose principle counts as a law of nature, rather than a gerrymandered variation on the law. If the alternatives are rivals in proleptic regularization, then we will prefer that which seems, in the principle it deploys, to offer the most accurate picture of the sort of pattern in question; in other words, that which approximates most closely to the proper account that we envisage.

Have we been provided with enough evidence to think that any broad-minded explanation, proleptic as it is, can be replaced by a narrow-minded alternative; specifically, by a decompositional alternative? Surely not. For all that we know so far, the *I**-state may be just the gerrymandered conditionalization of the *I*-state on the fact that *Cw*. It may be that the best picture of the causal pattern at work is projected in the *I*-account, so that only the *I*-state is the true explanatory kind. In that case we would not want to say that the account could be replaced by an *I**-story, since there would not be any genuinely explanatory *I**-state.

This prospect is threatening in particular because, for all that has been said, the *I*-state may be purely relational; it may resemble the state of an object which consists in its lying to the right of another. The narrow state defined by conditionalizing such a purely relational property could not be described in any other way and would be wholly parasitic on the first. It would be as insubstantial as the state of an object's being such that, given an appropriately placed second object, it lies to the right of the second.

We need to close the possibility that the *I**-state is insubstantial. There is no hope of arguing that the *I**-account can replace the *I*-story unless we do so; at best it will seem to be a misleading reformulation of that story. The argument of stages 2 and 3 will close the possibility and will establish the explanatory credentials of the *I**-account.

The argument of stages 2 and 3 will seem to be unnecessary on the following assumption: that it can always be established by analysis or whatever that the *I**-state is the same token and of the same type as some salient context-independent state. Thus the state of being such that, world willing, one knows that *p* might be analysed as the state of believing under certain stimuli that *p*. Similarly the state of being such that, world willing, one wants this cup might be analysed as the state of wanting the cup which one believes is presented in such and such a way. I assume that such heroic attempts at analysis are not guaranteed of success.[23] That being so, we must try another tack if we are to be sure of finding, for any broadly explained action *A*, a narrow-minded way of regularizing the behaviour.

Stage 2

The following two propositions are irresistible:

(a) Every action is at least partly constituted by a change in the context-independent character of the agent, usually a manifest change describable as a piece of behaviour.

(b) Every such change is caused by a narrow state of the agent, a state requiring nothing of the environment.

The first thesis is entirely uncontroversial, since all that it rules out

[23] A guarantee might seem to be provided by the thought that there is always something it is like to instantiate an *I**-state, and that under that presentation it is always independently plausible to link the state with the appropriate behaviour. But the thought is scarcely persuasive. For an exploration of related matters, see Simon Blackburn, *Spreading the Word* (Clarendon Press, Oxford, 1984), Chapter 9.

is action at a distance of the sort that would involve no narrow change in the agent. The idea of such action is empirically outlandish and not even clearly coherent. Suppose a table moves in the vicinity of a number of agents. Assuming that it moves by human action, in virtue of what does it move by one person's agency and not another's? In virtue of nothing, apparently, if the action need not involve any narrow change in the agent.

The change in question will usually consist in a discernible movement of the body, describable—in narrow terms—as a piece of behaviour. The obvious exceptions are the changes involved in mental acts like deciding or concluding. Even when these are broadly characterized—as in someone's wishing a demonstratively presented colleague well—the changes will not be overt. In what follows, however, I shall speak for convenience as if the narrow change required for every action was a piece of behaviour.

The second thesis can be established as follows. A narrow piece of behaviour could not be the unmediated effect of some event or condition independent of the agent; otherwise the behaviour would not be of the agent's making: he would be the arena in which it took place, not its author. The narrow behaviour must be traceable to a state of the agent, a state which may itself be narrow or broad. If it is narrow, then the second thesis is uncontroversially substantiated. But what if it is broad?

The broad state cannot be purely relational, since then the behaviour could once again be cast as the unmediated effect of some external event or condition; it must involve some context-independent structure in the agent. That structure may be thought to combine with the environmental correlate in the production of the narrow behaviour or to produce it on its own, whether or not mediating external influence. But it is difficult even to conceive of how the combination picture would work. And so in this case too the second thesis is substantiated: the narrow behaviour is caused by a narrow state of the agent.

Our two theses put us in a position to identify a narrow replacement explanation—that is, regularizing explanation—for the broad-minded account which invokes the *I*-state. We know that the *I*-state and the *A*-action have each got narrow counterparts. Let us ascribe the narrow behaviour involved in the action by '*Ba*'. And let us ascribe the narrow state which produces this by '*Na*'. We may assume that if *Ba* and *Cw*, then *Aa*. But if *Na* and *Cw*, then *Ba* and *Cw*. So if

Na and *Cw*, then *Aa*. That means, I shall take it, that reference to the fact that *Na* and *Cw* can regularize the fact that *Aa*.[24] We have therefore found a narrow replacement explanation for our original *I*-account.

Have we yet secured our goal? Have we shown that under a regularizing conception of action-explanation, the broad-minded explanation of an action is always dispensable, at least in principle, in favour of a narrow-minded one? Not quite. The trouble is that for all we have been told, the *N*-state need not have any claims to be a psychological one; it might be purely neurophysiological, as the letter '*N*' was meant to suggest. We will reach our goal only at the end of stage 3.

Stage 3

The stages of our argument so far are complementary. Stage 1 gave us a narrow psychological state which we could not be sure was substantial. Stage 2 has given us a narrow substantial state which we cannot be sure is psychological. Happily, the two results can be married so as to produce the offspring we want.

Let us return to the case where the *I*-state helps explain the *A*-action. Our conceptual decomposition had given us, as narrow counterparts, on the one hand the *I**-state, on the other the *A**-action. What the theses of stage 2 tell us is that in any such case there are also two independently describable narrow counterparts. For the broadly characterized action we have the narrow piece of behaviour, ascribed by '*Ba*'. For the broad intentional state we have the narrow state which causes that narrow behaviour; this is ascribed by '*Na*'.

What is the relationship between '*Ba*' and '*A*a*', '*Na*' and '*I*a*'? I shall argue that it is as follows. If *Ba*, then *A*a*; but not vice versa. And if *Na*, then *I*a*; but not vice versa.

We may assume that if *Ba* and *Cw*, then *Aa*; but not that only if *Ba* and *Cw*, then *Aa*: after all, a piece of behaviour not satisfying the *B*-description might serve equally well to ensure that *Aa*. By contrast we know, not just that if, but also that only if, *A*a* and *Cw*, then *Aa*.

[24] I commit myself to the principle implicit here while discussing a different issue in Graham Macdonald and Philip Pettit, *Semantics and Social Science* (Routledge and Kegan Paul, London, 1981), 125–6.

It follows that if *Ba*, *A*a*, since the fact that *Ba* ensures, given *Cw*, that *Aa*; and that ensures that *A*a*. On the other hand it does not follow that if *A*a*, *Ba*. It may be true that *A*a* other than through its being the case that *Ba*.

A similar argument applies to '*Na*' and '*I*a*'. We assume that if *Na* and *Cw*, then *Ia*; and we know that if and only if *I*a* and *Cw*, then *Ia*. We can deduce that if *Na*, then *I*a*; but not if *I*a*, then *Na*.

The relationships characterized fit with the supervenience picture mentioned in the last section. The *N*-state that we have identified is one on which the *I**-state supervenes; it is a state such that the *I**-state cannot cease to obtain without its also ceasing to obtain. Similarly the *B*-behaviour is something on which the *A**-action supervenes. In each case we have a truth expressed in one terminology—that of '*N*' and '*B*'—which ensures a truth expressed in another—that of '*I**' and '*A**'.

These supervenience relationships are sufficient to resolve our stage 1 worry. We know that the *I**-state is not insubstantial, since it supervenes on the undoubtedly substantial *N*-state. This means that we can think of the explanation which invokes the fact that *I*a* and *Cw* as revealing the antecedents of the *A*-action in a more illuminating way than the *I*-account. We can think of it as regularizing the action in terms of two factors rather than one, and as providing a better picture of the underlying causal pattern.

We have at last reached our goal. We can say that if action-explanation is regularizing in nature, then there will be a narrow-minded replacement available for any broad-minded story. This means that there will be little temptation to think of broad intentional states as genuinely mental. Unlike the narrow states, they will not be true explanatory kinds. It will be natural to cast them as hybrid states, straddling the psychological and the environmental. The elements of the hybrid *I*-state will answer to the fact that *I*a* on the one side, and the fact that *Cw* on the other.[25]

One final note. On the supervenience picture developed in this section, it will make for economy if we think of the *I**-state and the *N*-state as one and the same token state, albeit tokens of different types.

[25] One might take the view that broad types of state serve important explanatory roles outside the province of action-explanation, and view them, for that reason, as genuine explanatory kinds. See Kim Sterelny, 'Is Semantics Necessary? Stephen Stich's Case Against Belief', forthcoming in *Australasian Journal of Philosophy*.

In particular, it will enable us to see how, short of overdetermination, we can describe both the I^*-state and the N-state as the cause of the A^*-behaviour. Such an identity hypothesis is generally assumed in the literature.

5. *The Alternative Account of Action-explanation*

I come now to my alternative, non-standard account of action-explanation. We know that all action-explanation gives us a reason, a cause, and a pattern. What we wish to understand is why mention of the reason or cause, in a manner which provides evidence of a pattern, is indeed explanatory, enabling us to find the effect intelligible. Specifially, we need a story about this which does not turn on the regularizing impact of the explanation.

On the regularizing account, explanation is marked by the role in which it casts the principle that it embodies. That principle is taken to enunciate, or at least provide evidence for, how the world as a matter of fact works. It is a given such that if the explanandum can be paired with a cause and the cause–effect pair subsumed under the principle, then the explanandum is intelligible.

This observation is the cue for envisaging an alternative. It might be the case that the principle represented a norm at which the world or some part of the world aimed, rather than just a datum about how it worked. In that case the explanandum would be made intelligible, not by being shown to exemplify the world's regular mode of operation, but by being depicted as something that had to happen if the world was to continue to satisfy the principle that represents its norm. We can speak here, in a slightly artificial usage, of normalizing explanation. The principle is a norm in the light of which the explanandum is required, rather than just a datum in the light of which it is regular.

Suppose that we are given an event E, a cause C, and a principle to the effect that if a C-type event occurs, then so does an E-type one: for short, if C, then E. We might formulate an explanation of E by citing the principle, citing the cause, and concluding: therefore E. But this formulation in itself leaves it unclear whether the force of the explanation is regularizing or normalizing. The regularizing thought is: if the principle generally obtains, and the antecedent is satisfied in this case, then the consequent is fulfilled too. The normalizing counterpart is: if the principle is generally to obtain, and the antecedent is satisfied in this case, then the consequent has to be fulfilled too. Either thought might be expressed in the original deductive formulation.[26]

We look everywhere for regularizing explanations, since we believe that the world generally works to uniform patterns. We will look for normalizing accounts only in domains where we think that selection or design or whatever has been effective in ensuring that the systems there satisfy this or that norm. We will think of the systems as being constructed in such a way that the rule of basic, brute pattern means that at a higher level of characterization the systems answer to the appropriate norms.

These assumptions mean that normalizing explanations can never be taken to deploy principles which enunciate, as distinct from providing evidence for, fundamental underlying patterns. We saw that regularizing explanation may be proper or proleptic, depending on whether the principle deployed plays one or other of these roles. There is no ground for drawing a similar distinction among normalizing accounts.

In order to elaborate the notion of normalizing explanation, it will be useful to construct an illustration. I apologize for the science-fictional character of the example I have chosen. It helps me to avoid the complications which all real-life cases raise.

The Martians, omniscient and beneficent, have made a gift to the university. They have put a set of robots to work in the gardens, promising to monitor and maintain their performance. We look on in awe. Here on a lawn is one of the metallic creatures, busily moving with a probe about the surface. There beneath a tree, scratching and scraping among the leaves, is another. What on earth are they doing?

The explanatory desire behind this question is a wish to be able to substantiate our conception of the robots as gardeners; we have sufficient faith in the ingenuity and good will of the Martians to think that it applies. We want to be able to see how their various behaviours—other than those we come to regard as mere noise—contribute to the gardening task. We will scan the behaviour, trying in each case to view the initiative as a gardening response to a situation of need: in particular, a situation which it is plausible to think that the robot registers in some way. We will rotate the situation and

[26] Here I am grateful to the letter if not the spirit of some remarks in T. L. S. Sprigge, 'Final Causes', *Aristotelian Society Supplementary Volume* xlv (1971), 161. See also Andrew Woodfield, *Teleology* (CUP, Cambridge, 1976), 84 and 91.

behaviour beneath our gaze, looking for descriptions '*S*' and '*Gx*' such that the following principle is undeniable: if a system registers that *S*, and is bent on gardening, then, other things being equal, it *G*s.

We can easily envisage illustrations of this exercise. We ponder about the robot on the lawn until we conjecture that it registers that the grass is matted and, being a gardener, acts so as to aerate the roots. Again we puzzle over the creature beneath the tree until we realize that it probably registers that the leaves are decomposing and, in fulfilment of its gardening project, is doing what it can to rake them up.

The exercise is an example of normalizing explanation. Our conception of the gardener gives us a vague cluster of norms such that we expect any gardening agent to satisfy the antecedents of some, and, when it does, generally to satisfy the consequents too: generally rather than universally, because we can leave room for various blocks and imperfections, even ones we don't fully understand. We substantiate that conception when we see various robotic behaviours as occurring in order that such norms should be fulfilled.

The difference between regularizing and normalizing explanation comes in the assumptions against the background of which the explanation is offered. In the example on hand the assumption under which explanation is pursued is not just the regularizing thought that there must be some principles to which the domain of robots works. It is the stronger assumption that there are certain gardening principles to which, or at least to some of which, it must work in particular.

Suppose that we have a successful explanation of an event *E*—say, a piece of behaviour—by reference to an antecedent *C*—a registering of a situation of horticultural need. The idea is that we know more than the regularizing principle that if *C*, then *E*. We also know that it is because *E* is necessary for the satisfaction of that law—call it '*L*'—that it occurs. The assumptions are, on the one hand, *L*; on the other, if it is the case that if *L* then *E*, then *E*: this is another way of saying that if *E* is necessary for the satisfaction of the law, then it occurs.

The relation between these assumptions needs further clarification. The straight *L*-assumption is: if *C* then *E*. The other is that if it is the case that if *L* then *E*, then *E*. The consequent is the same in each case, so let us compare the antecedents. In the one case we have simply *C*; in the other, if *L* then *E*, or, more fully spelled out: if it is the case that if *C* then *E*, then *E*. The first antecedent ensures the truth of the second, as can readily be checked. And so the first, less complex assumption entails the truth of the other one.

But this seems to raise a problem. For it means, doesn't it, that whenever someone is possessed of an assumption sufficient for a regularizing explanation, he can also lay hands on an assumption sufficient for a normalizing one?

Normalizing explanation is defined so that this problem does not arise. The ground on which the normalizing assumption is believed cannot just be that the regularizing assumption holds. One must have a reason for thinking that if *E* is necessary for *L* then *E* occurs, over and beyond the knowledge that *L* obtains. This stipulation is not arbitrary, for it is fulfilled in the robotic example and in parallel cases. Given the preconception that the robots are gardeners, and given that *L* is a gardening norm whose antecedent they fulfil, we have a deep-seated reason for thinking that if *E* is necessary in a particular case for *L*—this, because the antecedent *C* obtains—then *E* occurs.

Generalizing our example, it seems that we are in a position to seek normalizing explanations whenever the following conditions hold. First, there is a more or less well-demarcated domain of explanatory concern, involving systems like the robots. Secondly, the systems in that domain are designed or selected or whatever so as more or less perfectly to fit a certain conception. And thirdly, we understand the conception sufficiently to be able to spell out norms at least some of which must apply to beings that fit the conception.

The most obvious case where these conditions are satisfied is with artefacts of a certain complexity. Given a chess-playing computer program, a pocket calculator, or even a central heating system, it is clear that there is room and need for normalizing explanation. The explanation will substantiate our preconception of the system, relating its behaviours to antecedent conditions in a manner which shows those behaviours to be the sort of thing required of a system of that kind.[27]

The conditions are also satisfied, I believe, in other circumstances. Take the case where we know that certain selectional pressures have

[27] The case of the chess-playing program will recall Daniel Dennett's notion of the Intentional stance. See his classic paper 'Intentional Systems', reprinted in *Brainstorms* (Harvester Press, Hassocks, 1978). The normalizing stance is a closely related idea.

governed the emergence of a species. With some skill and imagination, we may be able to work out a corresponding conception of the survivors, identifying norms which we think they must instantiate: norms such as those relating needs to efforts at fulfilment. This underpins a familiar sort of ethological normalization. The exercise is duplicated at another level of biology, when we concentrate on the selection of sub-systems and identify norms that we think these in turn must fulfil: say, the norms that require to be satisfied by a digestive system of a certain kind. At each level we elaborate a conception of the system in question, and we substantiate it in the identification of appropriate patterns of cause and effect.

Normalizing explanation should not be an entirely unfamiliar category. It is an instance of a widely endorsed conception of teleological explanation. According to that conception, an event E is teleologically explained as the means to a goal G if it is caused to occur by the circumstance of an E-type event's being necessary for the goal.[28] The conception fits our case, where the circumstance of an E-type event's being necessary for the obtaining of the C–E principle—a circumstance which is independently describable just as the occurrence of a C-type event—causes E to occur also.

So much by way of elucidating the notion of normalizing explanation. There are two further tasks for this section. First, I must show that action-explanation can be cast as an instance of normalization; and then, I must provide an argument that it ought to be seen in this way.

The three conditions sufficient for normalizing explanation are fulfilled in the action case. To begin with, we can certainly demarcate the class of human agents more or less adequately. We may have doubts about how far to include the immature and the demented, but in general we are quite clear about whether we are dealing with a proper human agent.

Secondly, we have reason to think that human agents will exhibit such actions, and indeed attitudes, as conform to a certain conception of what a normal person is. Every society requires as a condition of full membership that those who belong to it think and act in a manner which can be squared with such a conception, however

[28] This conception can be traced in Charles Taylor, *The Explanation of Behaviour* (Routledge and Kegan Paul, London, 1964); Larry Wright, *Teleological Explanation* (University of California Press, London, 1976); and G. A. Cohen, *Karl Marx's Theory of History* (Clarendon Press, Oxford, 1978).

loose. Social training is a process of moulding children so that they come to fit that ideal.

On this picture, training plays the role that design plays with an artefact like the chess-playing program or indeed the horticultural robot. To invoke it in this way, however, is not necessarily to endorse any sort of social relativism. We may still expect to be able to develop our conception of the human agent so as to encompass other cultures. One ground for that expectation is that no social shaping could produce agents who lived up to the local conception of the normal person unless evolution and biology provided appropriate predispositions. Those predispositions are the common inheritance of all members of the species, and suggest the possibility of intercultural accessibility.[29]

Thirdly, not only does social training ensure that people will fit a certain conception of the normal agent; it also enables each one of us to spell out norms which we expect such an agent to honour. He will satisfy the antecedents of a good many of the norms, and when he does so he will generally satisfy the consequents too. These norms include principles about the attitudes which people will form in the light of certain data or stimuli, and principles about the actions which they will choose in view of those attitudes. The principles fitting the belief–desire schema are normative expectations of precisely this kind; specifically, they are expectations governing the appearance of actions.

The fulfilment of our three conditions means that we can depict action-explanation as normalizing in nature. Under the picture on offer, we approach people in explanation with the presumption that the things they do—as distinct from reflexes, jerks, and the like—are prompted by beliefs and desires in such a way that they satisfy the appropriate principles. We presume that in general they are behaviourally rational, as indeed we make the presumption, in dealing with the formation of their beliefs and desires, that they are attitudinally so.[30]

This picture of action-explanation suggests that the exercise is

[29] For an a priori argument in favour of intercultural accessibility see Donald Davidson, 'On the Very Idea of a Conceptual Scheme', reprinted in *Inquiries into Truth and Interpretation* (Clarendon Press, Oxford, 1984).

[30] I have tried to characterize the distinction—and connection—between behavioural and attitudinal rationality in Macdonald and Pettit, *Semantics and Social Science*, 58–61.

rather like that which we imagined in dealing with our horticultural robots. It is an attempt to find such a characterization of people's doings as will enable us to substantiate our conception of them as more or less rational. This conception is more general in scope than that which we deployed with the robots, but it plays the same role. There we assumed that given the ends of a gardener, the robots were more or less rational subjects; we focused on action, but the assumption was also relevant to their attitudes, in particular to their dispositions to register and infer. Here we assume that given the rather more varied ends of the human being, ordinary agents are also more or less rational. Normalization in both cases comes to rationalization.

We have been considering the question whether action-explanation can be depicted as an instance of normalization, and, given that it can, the next issue is whether it ought to be cast in this way. Before we go to that question, however, there is an objection which must be turned aside. This is that on the normalizing picture, people are represented as selecting their actions in order to satisfy belief–desire principles, and that such a representation is not true to the phenomenology of deliberation.

The objection forces us to recognize that whereas the robots may have no conscious life, people certainly do. They don't act blindly in a manner that happens to sustain belief–desire principles. They act for reasons, and for reasons of which they can usually offer an account. The normalizing picture must be shown to fit with these facts about their subjectivity.

The claim is that no fit is available. It is suggested that under the normalizing picture, we must imagine the agent considering the sort of norms invoked in explanation and then acting so as generally to satisfy them. He would note the explanatory assumption that if someone believes this and desires that, then he acts in such and such a manner; he would recognize that he fulfils the antecedent of that assumption; and he would act so as to make the assumption true. Clearly, this is an absurd image of practical reasoning.

Happily, the image is not an essential part of the normalizing picture. It comes of a confusion between the third-person, explanatory version of a rational norm and its first-person, deliberative reading. Distinguish these and the objection is undermined.

The explanatory version of a belief–desire principle is roughly: *X* desires that *p*, believes that not *p*, and . . . , so he does such and such. The deliberative version, since it is appropriate to *X*'s own point of view, is rather of the form: it is desirable that *p*, it happens that not *p*, and . . . , so such and such needs doing. To say from the explanatory angle that *X* acts so as to satisfy the norm in question is to suggest that the deliberative version of the norm constrained him, consciously or unconsciously, not that the explanatory version did so. In its deliberative aspect, the fact that the norm applied constitutes a plausible reason for acting, and so the phenomenology of deliberation is preserved.

Let us agree that action-explanation, just as it can coherently be cast as regularizing in nature, can also be depicted as normalizing. The question to which we must now turn is whether it ought to be taken in this way. Assuming that one or the other account is the correct one, is there any evidence to suggest that action-explanation is better depicted as normalizing? I believe that there is.

The evidence I have in mind is the fact, already noted, that the principles engaged in action-explanation are a priori in character. It is not unthinkable that a priori principles should serve in a regularizing form of explanation. But if action-explanation is of the regularizing sort, the fact that the principles are a priori remains unexplained. By contrast, it becomes readily intelligible under a normalizing view of that explanation.

On the normalizing picture, the principles of action-explanation, relating beliefs and desires to behaviour, are norms which express our conception of the normal human agent. They are like the principles in which we might formulate our conception of a chess-playing program, or pocket calculator, or central heating system. Understand what it is to be a normal agent and it will be undeniable, for each of these propositions, that anyone who satisfies the antecedent will, other things being equal, satisfy the consequent too. That is to say, the principles will be a priori knowable.

That the principles are a priori does not mean that it is a priori that every normal agent satisfies the antecedents of all; nor that it is a priori that every noise and movement of every normal agent is subsumed by some principle; nor, above all, that every agent is normal. As we approach an agent with a view to substantiating our conception of the competent person, we must be prepared to find that he is not normal; that only in some of his responses does the conception

fit; and that it fits only so far as he satisfies the antecedents of a limited number of principles.

Neither does the fact that belief–desire principles are a priori mean that it is easy to apply them to actions. It is not as if we can postulate any old rationalizing set of attitudes for every piece of behaviour we want to explain. We will have to meet independent constraints on the ascription of attitudes, since the attitudes themselves will have to be more or less rationally explicable. And we will have to ascribe such attitudes as serve over time in relatively constant patterns of explanation.

That the principles are a priori means only that we hold on to them as fixed points when we embark on the explanation of action. We keep them out of the reach of questioning as we search around, subject to other constraints, for the principle which each action can be best seen as exemplifying. The principles are never themselves at risk of revision or rejection. They serve as standards with which every action has to be squared, not as generalizations which any action is likely to confound.[31]

To sum up then, the normalizing picture of action-explanation makes sense of the fact that belief–desire principles are a priori knowable, whereas that fact remains unexplained on the regularizing account. I conclude that not only can we see action-explanation as normalizing; if we are to reflect all the relevant facts, then we ought to see it in this way.

This completes the main business of the section, but before passing on I want to return to a crucial point of contrast between the two accounts of action-explanation. We saw that on the regularizing account, the intentional profile may command attention so far as it is a reason for the action—the assumption of rationality is a heuristic device—but it explains the action so far as it is a cause; specifically, so far as it is a cause which displays a regular connection with the effect. What we now have to see is that while the profile remains a cause of the action, the normalizing account has it explain as well

[31] That the principles of action-explanation operate like a priori standards I argue in Macdonald and Pettit, op. cit., pp. 93–101. For a critique see J. E. Tiles, 'Pettit on Revising our Understanding of Individuals', *Analysis*, xliii (1983), 189–93. I try to respond to that critique in 'A Priori Principles and Action-Explanation', *Analysis* (forthcoming).

as command attention in its role as reason; and, correspondingly, that it has the assumption of rationality play a constitutive as well as a heuristic part.

It should be clear that on the normalizing conception of action-explanation, the intentional profile remains a cause. The original argument still carries: namely that unless the attitudes involved were causally responsible for the appearance of the action, then they could not be invoked to explain it. At the weakest, the attitudes must have made the action more likely than it would have been in an otherwise identical situation where they were absent.[32]

But if the intentional profile is a cause of the action under the normalizing conception, it is not just because it has this status—or the status of a cause displaying a regular connection with the effect—that it explains. In order to explain it must point us, not to any old regularizing principle, but specifically to one on which the agent is targeted: a principle of a kind that he is constructed to satisfy. It does this so far as it is a reason. As a reason it points us to a belief–desire principle of a kind suited to figure in rationalizing explanation. It is clear, therefore, that the intentional profile has to be a reason, not just to attract notice, but also to explain.

Another way of putting this point is to say that under the normalizing conception, the assumption of rationality plays a constitutive as well as a heuristic role. To explain an action is to show that in the light of the agent's intentional states it was required by a norm, in particular by a principle of behavioural rationality. But this means that to explain an action is to show that in the light of those states it was the rational thing to do. Seeing an action as rational, then, is no longer just a way of coming to find the regularly connected causes that make it intelligible. It is what making the action intelligible consists in.

On the regularizing picture, action-explanation is a form of cause-giving explanation that may happen also to give us a reason. On the normalizing image, it is a kind of reason-giving explanation which

[32] The pattern of cause and effect is not enunciated by the normalizing principle; the principle merely provides evidence of its existence. It will be enunciated, if it can be enunciated, by a very different sort of law. The law will not be hedged by an open *ceteris paribus* clause and, if the principle relates broad phenomena, then the law will distinguish the context-independent and environmental components in antecedent and consequent.

happens, albeit not just incidentally, to give us a cause. The contrast between the two conceptions is nearly complete.[33]

6. *The Alternative Account does not make Broad-minded Explanation Dispensable*

If action-explanation were of the regularizing kind then broad states, as we have seen, would not strictly be needed in psychology. Broad-minded explanation would be in principle dispensable. We now know that action-explanation is normalizing in character, not regularizing. The question, then, is whether this saves broad states from the redundancy to which they are condemned under the other dispensation. I shall argue that it does.

We know that for any broad state ascribed by '*Ia*', there are an indefinite number of decompositional alternatives. We identified the *I**-state of being such that, if *Cw*, then *Ia*; and the *I*(@)-state of being such that, if 1 + 1 = 2, then *Ia*. These alternatives are designed so that they can apparently be invoked to explain anything which the fact that *Ia* explains: instead of this fact, we would call upon the fact that *I*a* and *Cw*, or the fact that *I*(@,*a* and 1 + 1 = 2, or whatever.

In discussing regularizing explanation, we agreed that with such alternatives only one type of state can be regarded as the true explanatory kind; the others must be seen as gerrymandered entities. Which state is to be preferred? In general, the type that is called upon in the explanation that best reflects nature's mechanical patterns. This line of thought led us to give preference, in any case of the sort envisaged, to the *I**-state rather than the *I*-state, or indeed the *I*(@)-state. The regularization which best seemed to mirror the pattern that a proper explanation would enunciate was that which called upon the fact that *I*a* and *Cw*.

Turning now to normalizing explanation, we have to see whether the same result falls out. We assume that the point of calling on the fact that *Ia*, or on any alleged alternative, is to normalize the corresponding behaviour. Well then, the first question is whether the old criterion of explanation-preference still applies, for if it does, then it will again give the plaudits to the *I**-state. I shall argue that it does not apply, and then I shall take up the further question of what substitute criterion is relevant in adjudicating between alternative normalizing accounts.

The guiding interest of regularizing explanation is to present the event explained as an unsurprising moment in the causal unfolding of the natural world. If there are two versions of such an explanation and one gives us a better picture of the mechanics involved, then it is inevitable that we shall prefer this account. The same does not hold of normalizing explanation, because its guiding interest is quite different.

Its goal is to present the event explained as something which was required of the relevant system, given a certain conception of that system; specifically, given a conception motivated by our beliefs about its design or selection or whatever. Suppose that we have two versions of such an explanation, and that one has the merit, perhaps at some cost in other regards, of answering more precisely to the mechanical structure underlying the system's performance. Will we favour this alternative?

Not necessarily. The alleged merit is quite extrinsic to the normalizing drive, for it does not mean that the event explained is any the more readily seen as normative for the system. It will not motivate a choice of the marked alternative, though it might serve to break a tie. And certainly it will not weigh against any cost that is intrinsic to the normalizing goal.

The point is best made by returning again to our gardening robots. As we survey the robots and wonder about them, we can clearly distinguish between two different explanatory desires. On the one hand we will want a normalizing account of their behaviour which shows in what respects they are gardeners. On the other, at least if we have any curiosity, we will want to understand how they

[33] It will be obvious from these remarks how close I am in spirit to those who have argued, incorrectly as it may be, that reasons are not causes. The tradition in question is characterized—and criticized—in Macdonald and Pettit, op. cit., pp. 80–93. I also find much that is congenial in the writings of Donald Davidson and Daniel Dennett on action-explanation, though I remain unsure in the case of each as to how far they resist the regularizing model. See Davidson, 'Actions, Reasons, and Causes', and 'Mental Events', and Dennett, 'Intentional Systems'. See also John McDowell, 'Physicalism and Primitive Denotation: Field on Tarski', *Erkenntnis*, xiii (1978). Many recent writers recognize that a cause-giving explanation need not be causal in the regularizing sense. See for example Christopher Peacocke, 'Demonstrative Thought and Psychological Explanation', *Synthese*, xlix (1981), 213, and Peter Achinstein, *The Nature of Explanation* (Clarendon Press, Oxford, 1983), Chapter 6. John McDowell emphasizes the point in 'Functionalism and Anomalous Monism', in Ernest LePore and Brian McLaughlin (eds.), *The Philosophy of Donald Davidson: Perspectives on Actions and Events* (Blackwell, Oxford, 1986); he thinks it is implicit in Davidson.

work mechanically: how they are constructed so that, exploiting the laws of nature, they succeed in sustaining their gardening profile. This is a desire for a certain sort of regularizing explanation.

Suppose now that we pursue the normalizing goal, and that, in explaining a particular piece of behaviour, we have to choose between two decompositional alternatives. Why is that robot scraping among those leaves? Because it registers that those leaves are decomposing and, being a gardener, does what it can to rake them up. Alternatively: because it is such that, world willing, it registers that those leaves are decomposing and, being a gardener—better perhaps, a gardener*—it acts so that, world willing, it does what it can to rake them up; and the world is willing, so it does what it can to rake them up. Of these roughly formulated alternatives, let us agree that the second more accurately reflects the mechanics involved. Ought we necessarily to endorse it, then?

Certainly not. The preference for such a parasitic alternative must strike us as bizarre. It will seem to spring from a confusion between the two different explanatory goals that we distinguished. When we are normalizing, why should we worry about the regularizing account that will be required once we turn our minds to the other task? The two tasks can be pursued separately. Running them together only introduces a further risk: namely that we will subject one or both to irrelevant criteria of achievement.

I conclude that the criterion of explanation-choice that operates with regularizing accounts ought not to be allowed to govern the selection of one normalizing story over another. Specifically, it ought not to be allowed to determine the judgement between decompositional alternatives. This result is important, for it means that the argument developed in section 4 is irrelevant, given that action-explanation is normalizing rather than regularizing. The argument gives us reason to prefer an *I**-account to an *I*-account if the aim is to regularize action, but not if the aim is to normalize it.

We turn now to the second question advertised. Given some decompositional alternatives for the normalization of an event, how ought we to judge between them? If the criterion invoked for judging between regularizing alternatives does not apply, then what yardstick takes its place?

The appropriate criterion is identified by reference to the guiding interest of normalization. The goal in normalizing an event is to substantiate a motivated conception of the system that occasions it.

Well then, if we are presented with decompositional alternatives, we will naturally prefer that which substantiates the conception most straightforwardly.

Let us return to the robot case, where we are choosing between different accounts of why the robot scrapes among the leaves. We already have two alternatives on hand: the *I*-account, as we may call it, and the *I**-account. We may add an *I*(@)-story as a third alternative. This would account for the event by reference to the fact that the robot is such that, if $1 + 1 = 2$, then he believes that those leaves are decomposing, and $1 + 1 = 2$. Of these accounts it is clear that the *I*-explanation substantiates the gardener conception of the robots most straightforwardly, and that we would naturally prefer this to the alternatives.

To resort to shorthand, our conception gives us immediate reason to think that if *Ia* then, other things being equal, *Aa*: if the robot registers that those leaves are decomposing, then it does what it can to rake them up. It gives us reason only indirectly to think that if *I*(@)a and $1 + 1 = 2$ then *Aa*, for in order to see this we must do some ratiocination. And equally it offers us only an indirect reason for believing that if *I***a* and *Cw* then *Aa*. The *I*-account is salient to someone who employs the conception, the others are not.

The reason why the *I*-account has this different status is that our conception of a gardening robot has a broad character. It offers us the prospect of a system which achieves a certain environmental equilibrium: in various conditions defined in terms of the system's surrounds, appropriate equilibrating responses are forthcoming. Given that conception, it is not surprising that the antecedents of the norms it embodies are broad states of the system, and indeed states which can be identified without recourse to the fact that $1 + 1 = 2$.

The fact that the system satisfies a norm with a broad antecedent and broad consequent means that it must be constructed so that when the corresponding context-independent condition is realized, it evinces a suitable context-independent response. But this does nothing to motivate a preference for the *I**-account over the *I*-account. Our focus in deploying the gardener conception of the robot is not on the isolated system. It is on the system considered in environmental integration.

We were dealing with the question which criterion ought to govern the decision between decompositional alternatives for normalizing an event. We have seen that the judgement has to be made on the

basis of which candidate answers most directly to the conception deployed in the normalization. With this criterion in hand, let us now ask how the decision must go between broad-minded and narrow-minded accounts of human behaviour.

What is true of the gardening robot is true also of the human agent. To return to our original example, suppose that '*Ia*' stands for '*a* wants this cup' and '*Aa*' for '*a* grasps this cup'. Our ordinary broad conception of agents allows us to see directly that in appropriate circumstances the intentional state requires the action. But this is not so for the connection between the corresponding *I**-state and the action. Some inference is required to see that if the agent is such that, world willing, he wants this cup, and the world is willing, then he grasps this cup. The principle does not fall so directly out of the conception.

I conclude that if action-explanation is normalizing, as I have argued that it is, then the judgement between such decompositional alternatives as the *I*-account and the *I**-account goes the other way from how it would if action-explanation were regularizing in character. This means that in such cases we should regard the *I*-type of state as the true explanatory kind and the *I**-type as something gerrymandered out of it. More particularly, it means that we must regard the broad-minded explanation as indispensable. We may be able to reformulate a normalization that calls upon the fact that *Ia* so that it invokes the fact that *I*a* and *Cw*. But the reformulation is misleading, since it misdirects us about the explanatory kind involved. And certainly it does not provide us with a replacement for the original account.

The considerations marshalled so far in this section concern normalization in general. I believe that they are sufficient to show that if broad intentional states are invoked to normalize action, then we may regard them as genuinely mental states, not mere hybrids. But there is also a special argument that can be made for that result, and I would like to conclude the section by mentioning this.[34]

The special argument is best presented as a response to a special objection. The objection is that, whatever of normalizing concerns in general, our interest in psychological explanation is in providing an

[34] The argument connects closely with Gregory McCulloch's paper in this volume.

account of how things get organized within the head, or at least within the hide, of the subject. We may operate with a broad conception of the normal agent, but nevertheless, so the line goes, we aspire to a story about context-independent connections. That being the direction of our interest, the lesson drawn is that we should see every broad-minded normalization as an account that goes proxy, in the absence of a better understanding, for a purely narrow-minded story. We should view every *I*-account as a story that is better expressed in the form of an *I**-explanation, and we should take the *I**-state as the true explanatory kind.

This objection foists a cognitive psychological ambition on folk psychology, as it is called. The cognitive psychologist, on at least one reading of his strategy, argues as follows. He admits, or at least may admit, that on the broadest and ultimate front the human agent is an environmentally integrated system requiring broad-minded normalization. But he argues that on a nearer and narrower front each agent must be a system designed to satisfy corresponding context-independent norms. The cognitive psychologist seeks to work out a conception of such a narrow system, and perhaps of narrower and narrower sub-systems, with a view ultimately to seeing how neurophysiology supports it. The objection suggests that this is the enterprise in which we folk psychologists are already unwittingly engaged, and that for that reason our broad-minded normalizations should be seen as our best attempts to get at hide-bound connections.

I have nothing against the idea of cognitive psychology. On the contrary, I believe that the top-down strategy just sketched is entirely plausible. What it promises, in my view, is the possibility of normalizing human performance at lower and lower levels. Leaving the broad-minded conception of the agent behind, the cognitive psychologist tries to work out a conception of the sort of narrow system—and ultimately systems—that would subserve the broader complex, and he then seeks to substantiate that conception in the identification of functional elements such as those responsible for information processing, information storage, information retrieval, and the like. The fact that the agent can be submitted to successively lower levels of normalization is in no way inconsistent with my claim that the topmost level, namely that of broad-minded explanation, is not dispensable in favour of narrower accounts. It remains

indispensable, because it is alone in substantiating its particular conception of the human person.[352]

What is inconsistent with my claim, however, and it is the core of the objection presented, is the view that folk psychology is a vulgar version of the scientific enterprise: that it is really proto-cognitive psychology. Admit that view and the broad-minded character of ordinary action-explanation will seem to be of no consequence. It will be explained away as a product of the rough and ready nature of folk concepts.

As against this objection, we might argue that the depiction of ordinary action-explanation as proto-science is extremely uncompelling. I prefer to rely, however, on what I have described as my special argument for the indispensability of broad-minded explanation. The argument is that the conception of the human agent deployed and substantiated in such explanation is of outstanding intrinsic merit, and that the exercise cannot be seen, therefore, as something that goes proxy for a context-independent level of normalization: that is, for the deployment and substantiation of a distinct narrow conception of the person.

In order to motivate this thesis, let us go back once more to our robotic gardeners. For all that we have postulated so far, they are blind reasoners: creatures which, without engaging in any deliberation, let alone in any process of review and criticism, manage to act in a rationalizable fashion. Imagine now that not only do they conform to gardening norms: they do so by understanding those norms and, being committed to their implementation, by monitoring themselves and one another for their degree of norm-fulfilling success.

Even under the less complicated picture of the robots, it is plausible that we should want to substantiate the gardener conception of them, whether or not this is a first step in exploring their cognitive psyche; after all, their ultimate telos is to be gardeners. But under the picture now projected, we have reason multiplied to be attached to the substantiation of this broad explanatory conception. It is only by deploying that conception that we can hope to be able to understand the robots as they understand themselves and one another. It is only by substantiating that conception that we can participate with them in a shared understanding of what they do and why they do it.

[35] This means that I can endorse the substance of what William G. Lycan describes as homuncular functionalism. See his 'Toward a Homuncular Theory of Believing', *Cognition and Brain Theory*, iv (1981).

The lesson of the robotic parable is obvious. The norms to which each of us feels himself bound, the norms that we invoke in deliberation and review, are those associated with the broad-minded conception. Equally the norms to which we expect one another to be faithful are principles which involve environmental requirements. When any one of us, therefore, plays psychologist with his fellows, he will have reason to take broad-minded explanation seriously. He may have a particular commitment to developing cognitive psychology, but, even so, he cannot think that ordinary action-explanation is uninteresting.

The fact is that only by deploying and substantiating the broad conception of agents can any one of us hope to be able to understand his subjects as they understand themselves. It is only by cleaving to the broad-minded way that he can aspire to psychological understanding, as we might say, of a participant character. Such an understanding is obviously attractive in its own right, being a condition of interpersonal contact. But it also must appeal for instrumental reasons. It is essential, for example, if someone wishes to influence his fellows in the distinctively human manner, making it clear where the actions they evince or the attitudes they form are based on oversights and mistakes.

In the main part of this section we saw that the criteria relevant to normalizing explanation would suggest that the true explanatory kinds invoked in action-accounting are the broad *I*-states, not their *I**-surrogates. The special objection which we have been considering concedes that the argument offered may work in parallel cases but claims that it fails in the psychological one. The reason alleged is that the broad conception deployed in action-explanation is of no intrinsic interest, serving only to adumbrate the narrow conception which would suit a cognitive psychology. The special argument rebuts this charge, for it shows that on the contrary the broad conception is essential to the achievement of participant psychological understanding.

In conclusion, a comment on terminology. It is common nowadays to counterpoint the notions of folk psychology and scientific, in particular cognitive, psychology. I think that the contrast is misleading, for the category of folk psychology is too mixed a bag. It includes not only our deeply embedded ideas about how action should be explained, but also every currently fashionable story about the nature of dreams, or memory, or drives, or whatever. The special

argument which we have just considered suggests that we ought to replace the contrast by one which sets participant psychology on the one side, observer psychology on the other. Participant psychology is what we practise when we seek to rationalize human actions—and of course attitudes—displaying them as the product of a greater or lesser fidelity to appropriate norms. Observer psychology is everything else. It probably attains its highest form in cognitive psychology, but it also includes the shifting conjectures of less scientific folk. The contrast between folk and scientific psychology was drawn on grounds of precision. The contrast between participant and observer psychology is based on grounds of purpose.

7. *Conclusion: The Standard and Alternative Views of Intentional States*

The more or less standard, functionalist view of intentional states depicts them as causal roles or transitional susceptibilities. To instantiate such a role or susceptibility is to be such that, given certain inputs, and given other interlocking intentional states, certain outputs are ensured. To believe that p or q is to be such that, given evidence that not q, and given the appropriate logical beliefs or dispositions, one moves to the belief that p. To believe that p is to be such that, given the desire that not p, one acts so as to change the status quo. And so on.

There are many issues about the precise interpretation of functionalism. I shall ignore all but one. This is a question about how to identify an intentional state functionally when there are a number of functional candidates available.

Suppose that we have a broad functional state S, a state which is characterized in part by reference to broad inputs and broad outputs. To instantiate it is to be such that, among other things, if C and . . . , then E, where C and E are events which require the world to be thus and so. Consider now the narrow functional state which S presupposes. This will be a state characterized solely in terms of narrow inputs and outputs; specifically, it will be a state such that fulfilment of it in a congenial environment ensures the instantiation of S. We can identify it as the state S^* of being such that, among other things, if C^* and . . . , then E^*, where these events respectively guarantee C and E in a cooperative world.

Grant that there is an intentional state of the bearer with which S and S^* are associated. The question then is; which functional state is

the intentional one? They are different states, having different existence conditions, so they cannot both be the intentional state. We must decide between possibilities like the following. (1) S is the intentional state, and S^* is merely a precondition of it. (2) S^* is the intentional state, and S is a hybrid construction out of it. (3) S is the intentional state, but its only psychologically significant component is S^*.

Within the functionalist literature the general preference is for a position like (2) or (3). The reasoning is that the intentional state is a causal role, and that the state which does causal work, so to speak, is S^* rather than S. The environmental condition which makes the difference between S and S^*, after all, is acknowledged on all sides to be causally inert.[36] If S^* is taken to be suitably intentional, (2) will seem attractive; if not, (3).

This reasoning might be paralleled by another argument. Suppose we take intentional states to be causal roles in the sense of causal–explanatory roles. To instantiate an intentional state, then, is to be such that, given certain inputs and certain collateral states, such and such outputs are causally explainable. Suppose further that we take causal explanation to be regularizing in character. It will then follow that position (2) or (3) is the most reasonable one. All that can be causally regularized by S can be regularized by S^*, so parsimony and indeed simplicity will suggest that we should identify the intentional state, or at least its psychologically significant component, with S^*.

So much for the orthodox answer to the question raised. I am now in a position to indicate the view of intentional states to which my approach naturally leads. It is distinguished by the fact of endorsing the first possibility mentioned, providing an unorthodox answer to our question.

I am happy to go along with the picture of intentional states as causal roles, so long as this means causal–explanatory roles. What I wish to insist on, however, is that the sort of causal—better perhaps, cause-giving—explanation which is relevant to the identification of those roles is the normalizing, in particular the rationalizing, variety. I have argued that intentional states explain by rationalizing, and I have assumed that they are themselves explained in a similar manner.

This is to say that intentional states are rationalizing roles rather

[36] See n. 18.

than—as in the explanatory version of standard functionalism—regularizing ones. They are profiles characterized by the fact that given such and such inputs and collateral states, they render certain outputs rationally necessary. Those outputs will include transitions to other intentional states as well as initiatives of action.

With this conception of intentional states, it is clear how we should approach the question whether the intentional state in our example is S or S^*. We must conceive of S and S^* as rationalizing roles, and we will want to know which plays the role corresponding to the rationalizing potential of the state under discussion. In all the examples imaginable, the answer will be the broad state S. This will be cast as the intentional state, and the narrow counterpart S^* will be seen as a precondition of it. That is to say, position (1) will command allegiance; the unorthodox answer will triumph.

On the view of intentional states to which I am committed, then, they are states that become salient and indeed visible only from the standpoint of rationalizing explanation. Adopt the assumption of rationality *vis à vis* agents and their intentional states are promptly highlighted. Shift stance to consider agents in a regularizing mode, or even at some lower level of normalization, and these profiles will vanish from sight. They are real attributes of people, but, like secondary properties, they manifest themselves only to observers who occupy a particular perspective: specifically, the perspective of a certain explanatory disposition. They are rational-explanatory kinds, and they become salient only from a rational-explanatory viewpoint.

The intentional states that we ascribe to ourselves and one another are the precipitates of our explanatory practice. Cast action-explanation and attitude-explanation as regularizing and those states will come out in the standard, functionalist mould. Cast it as rationalizing and they will take the shape that I have described. Here as before we see the importance of the issue about explanation. There is probably no deeper question to be found in the province of philosophical psychology.

CHAPTER 2

SCIENTISM, MIND, AND MEANING*

GREGORY McCULLOCH

Introduction

This paper falls into three parts. In the first I describe how the classical notion of meaning is intimately bound up with the nature of folk psychology—the 'discipline' in terms of which we understand ourselves and each other as ordinary thinking (etc.) beings. I set out the view that in giving an account of what folk psychology is one works with some classical notion of meaning, by which I mean the reference-determining sort of meaning familiar from the work of Frege and his more recent followers. In this sense, folk psychology and classical meaning seem to stand or fall together, since it is hard to see what other purpose classical meaning is needed for. The emphasis of this first part is more on description than argument. I do not so much defend the view that there is an essential link between folk psychology and classical meaning as describe a link which is, I believe, generally accepted to exist. The business of supplying arguments occurs mostly in Parts 2 and 3.

The second part deals with the status of folk psychology as an account of what mental entities are really like. It is considered reasonable to expect that, one day, it will be possible to give explanatory accounts of something like behaviour that operate exclusively with predicates that are *solipsistic*, in a sense that I shall briefly explain. But this raises the question of the relationship between the discipline that would employ these predicates, and folk psychology

 I read an earlier version of this paper at the University of Nottingham in October 1983, and benefited greatly from the discussion that followed. I am particularly grateful to Bob Kirk, Nick Measor, and especially Robert Black. In Parts 2 and 3 I am indebted to the work of John McDowell: see especially 'Meaning, Communication, and Knowledge', in Zak van Straaten (ed.), *Philosophical Subjects* (Clarendon Press, Oxford, 1980), and 'Anti-realism and the Epistemology of Understanding', in H. Parret and J. Bouveresse (eds.), *Meaning and Understanding* (De Gruyter, Berlin, 1981). Finally, various changes for the better have been prompted by the responses of the editors to an earlier draft, and my thanks are due to them for this.