

Two kinds of purposive action

s.butterfill at warick.ac.uk

Abstract It is normally assumed that there is only one kind of purposive action. This article argues that there are two kinds of purposive action, which require different models of explanation. One kind of action is done without awareness of reasons; another kind of action is done because the agent is aware of reasons for that action. The argument starts by noting that philosophers disagree about what explains action. Some claim that actions are explained by impersonal facts, such as facts about how things should be or have been historically (e.g. Millikan, Stout). Others claim that actions are explained by mental states, such as beliefs and desires (e.g. Davidson, Velleman). These philosophers are usually regarded as offering conflicting accounts of one thing. However, they are best understood as giving accounts of different models of action-explanation. Neither model fits every case, so there are at least two kinds of purposive action.

[10450 words, and 2600 words in endnotes.]

There is a standard distinction between response behaviour and genuinely purposive action. This is based on a further distinction between the kinds of explanation appropriate to each kind of behaviour. Response behaviours can be explained as governed by associations, whereas genuinely purposive actions require rationalising explanations. This is because genuinely purposive actions normally happen only when they fulfil an agent's purpose.

The category of response behaviour is, at first sight, heterogeneous. It includes reflex actions, learnt responses to stimuli and certain kinds of instrumental action (Dickinson 1994). Perhaps there is a single theoretical model which unifies the various types of response behaviours, but it is not obvious *a priori* that anything unifies them. Particular theories about the explanation of response behaviours are tested in laboratories. Contrast purposive action. It is generally assumed without

appeal to evidence that purposive behaviours are all of a kind, and that there is just one kind of rationalising explanation by which purposive behaviour is explained.

I'm going to argue for a distinction between two kinds of purposive action, *information-based but unreflective actions* and *actions done for reasons*. As a rough initial characterisation, information-based but unreflective actions are actions that occur because they are *in fact* reasonable, whereas actions done for reasons are actions that occur because the agents *take* their actions to be reasonable. Information-based but unreflective actions are reasonable and intelligent ways of fulfilling a purpose. There are reasons for which these actions occur, but the agent is not necessarily aware of having reasons for his action. The apparent concern with rationality is entirely nature's doing: the agent himself need not even be capable of caring whether his actions are reasonable. All the agent cares about is the outcome. In the case of actions done for reasons, on the other hand, the agent is in a position to care about the rationality of his actions, and not just about whether they fulfil his purposes. The agent is aware of his reasons for acting *as reasons*, and this awareness is partly what causes him to act.

By saying that someone who acts for a reason *cares* about the rationality of his action, I don't mean to imply that one of his goals is to act rationally, as if being rational were on a par with getting fed or paid. This concern with rationality is like an honest person's care for the integrity of his actions. Acting with integrity is not one among several goals, and neither need it be a means to an end. Instead, the concern to act with integrity can manifest itself in the way someone will pursue particular goals. There is nothing specially intellectual or theoretical in caring about integrity or rationality in this sense. Equally, by saying that someone who acts for a reason is *aware* of having reasons to act, I don't mean to suggest that reasons are necessarily foremost in his mind. In acting for reasons, we characteristically feel what we are doing to be reasonable, and we know we could explain ourselves if we had to. This is an important part of how it feels to act. That we often take for granted awareness of our reasons for acting is revealed by experiences in which we find ourselves acting without quite knowing why. It can be quite disturbing to suddenly have to stop and ask, 'Why did I do that?' Imagine, for example, that Ayesha surprises herself by kissing Tom in order to console him. In one sense, she

knows exactly why she is kissing Tom, for there is no doubt about her purpose. But, in another sense, she doesn't understand herself because she doesn't at first know why her kissing Tom should console him – is it the softness of her lips, or the solidarity demonstrated by a kiss, or what? It just instinctively came to her that kissing Tom would console him. If Ayesha can explain why she kissed Tom, it is only by means of the kind of narrative reconstruction that she would use to explain another's actions. Having worked out why kissing should console Tom, she carries on kissing him; but now Ayesha is conscious of her reasons for acting. Now her actions are done not just because they are in fact reasonable, but also because she finds them reasonable. And the explanation for her action is immediately apparent to her in the sense that she could explain herself without recourse to narrative reconstruction.

Not everything we do involves a concern with our reasons for acting, and nor should it. Some purposive actions are consequences of reasons that we are not aware of as reasons. Imagine we register the imminent departure of the last bus and start running. The running is purposive and there are good reasons for it, but we might be unaware of our reasons for doing it. Suppose, for example, that we miss the bus by a wide margin. Were we running because we thought this was a way of getting to the bus before it left, or because we thought the bus driver might see us running and wait for us? Maybe we were aware of particular considerations as our reasons for running, but it is more likely that we felt only dismay at the thought of missing the bus and weren't at all concerned about why we should run. Even so, there could be particular reasons why we ran for the bus, in the absence of which we would not have run for the bus. Actions such as running for a bus are sometimes information-based but unreflective: they are clearly purposive, but can be done without concern about, or awareness of, our reasons for doing them.

This intuitive characterisation of the notions of acting for a reason and information-based but unreflective action is, I admit, unconvincing as it stands. Most people would probably agree that we are sometimes but not always aware of which reasons we are acting for, but disagree that actions done in part because we are aware of our reasons for them are different in kind from actions done merely because we care about getting results. The purpose of this paper is to argue that there is such a

difference between kinds of action, and that some of our actions are information-based but unreflective while others are done for reasons. In the first part, I'll describe a theoretical way of characterising these two notions of purposive action. Then, in the second part, I'll suggest some empirical motivation for the distinction between them.

Before launching into the argument, why does it matter whether there is more than one kind of purposive action? Purposive action is relatively undemanding in terms of the cognitive capacities it requires. To act purposively is just to fit your actions to a purpose in whatever way you can. Infants, dogs and rats all do it. None of these show any sign of being able to reflect on reasons. They act intelligently; they are responsive to reasons, but never knowingly. They know how to act, but why they should act doesn't worry them. Unlike them, we live in a world of reasons, where actions are intelligible in a way that rats and dogs cannot even begin to comprehend. We are always trying to justify actions, to uncover reasons for them so that we can understand the story they tell. Our concern with reasons can't be entirely theoretical. It must make a difference to our actions. But what kind of difference does it make? Some philosophers think that our concern with reasons doesn't make any difference to the nature of simple things we do: they think that awareness of reasons only matters when we have more complex purposes than infants and dogs can manage. I disagree with this view, because we're not concerned about reasons only when they're complex. Often it is the simplest things that hold our interest, like a kiss or a smile. When the reasons get complex, even our understanding of our own reasons tends to lag behind our actions. When we put a concern with reasons into action, it is often only to do things for unsophisticated reasons and with simple objectives. We could easily have reasons and objectives like these quite unreflectively. So if there were only one kind of purposive action, what we do would be unaffected by whether or not we do it with an eye to the reasons for it. Our concern with reasons would be literally of no consequence. To understand how being concerned about reasons for acting makes a difference, we therefore need to understand how the actions this concern results in differ from unreflective actions that appear to be done with the same objectives and for the same reasons.

§1 Is the distinction theoretically viable?

How can we decide whether there really is a distinction between two kinds of purposive action? I think that a distinction between two kinds of action should be made 'in terms of the type of explanation which is to be given for the occurrence in question' (Brewer 1993, 274). For instance, we can be sure that response behaviour and purposive action are different kinds of action because there is a well-established distinction between corresponding kinds of explanation: associationist explanations for response behaviour, and rationalising explanations for purposive action.¹ Similarly, giving a theoretical characterisation of two kinds of purposive action requires distinguishing two different types of explanation. Whereas philosophers generally assume that there is just one kind of rationalising explanation for purposive action, I will argue that we should distinguish two different kinds of rationalising explanation. In fact, I think that two groups of philosophers – who I'll call *Teleologists* and *Mentalists* – have already offered accounts of the different kinds of explanation, although they were presented as different accounts of one kind of explanation.

1. The Teleologist and the Mentalist can both agree that whenever purposive actions are the result of an agent's encountering information, this is in part because he has acquired beliefs (or some kind of informational-states) relating to that information. Otherwise it would be a miracle that encountering information can facilitate action even after the original source of information has disappeared. The Teleologist and Mentalist can also agree that beliefs are causal factors in action. Despite agreement on these points, the Teleologist and Mentalist give different accounts of the role of beliefs in rationalising explanations of actions. This disagreement would be insignificant if the point of a rationalising explanation were only to say why an agent *actually* acts. However, the Teleologist and Mentalist both hold that a rationalising explanation characteristically reveals why an agent *should* act.² They agree that an agent would not act in the way he does were it not for his beliefs, but they disagree about how the fact that an agent has particular beliefs is relevant to explaining why he should act in the way he does.

To illustrate the disagreement between Teleologist and Mentalist, suppose that Ashwin buys lilies for Sarah. We can begin to explain why he did this by citing his

purpose: he intended to cheer Sarah up. The explanation is incomplete, however, because we need to explain why buying lilies is a rational way for Ashwin to achieve his purpose. Given his purpose, there are many possible considerations in favour of Ashwin's action. Buying lilies might cheer Sarah up because she adores them; or maybe she hates lilies but appreciates being given flowers anyway; or maybe the lilies have special significance for her. Any of these considerations could justify buying lilies for Sarah in order to cheer her up: impersonally speaking, each is a reason for Ashwin to buy lilies. However, not all of these considerations are necessarily relevant to explaining why Ashwin acts because they are not necessarily among Ashwin's reasons for buying lilies. Certainly, considerations of which Ashwin is completely unaware can't explain his action. It is clear, then, that agents don't always act on every available consideration in favour of their action. So we can ask what would make it the case that one rather than the other of the following possible justifications of Ashwin's action explains why it happens:

Justification (a):

Ashwin intends to cheer Sarah up.
Sarah likes lilies.

Therefore, Ashwin should buy Sarah
lilies.

Justification (b):

Ashwin intends to cheer Sarah up.
Sarah appreciates being given
flowers.

Therefore, Ashwin should buy Sarah
lilies.

A way of justifying Ashwin's action will explain why it happens only if it justifies his action in terms of the reasons for which Ashwin acts. So the question is this: Among the possible considerations in favour of Ashwin's action, what determines which are the reasons why he acts? In other words, what would have had to be different in order for Ashwin to have acted for reasons other than his actual ones?³ The fundamental difference between the Teleologist and the Mentalist is their conflicting answers to this question.

I'll look at the Mentalist first. The Mentalist thinks that only facts about the contents of Ashwin's state of mind can determine which considerations are reasons for his action. That is, a consideration is a reason for Ashwin's action in virtue of being the content of a mental state on which he acts. For example, the Mentalist

might say that Ashwin's acting on certain beliefs makes it the case that the contents of these beliefs are his reasons for acting and therefore explain why he should act. So the Mentalist holds that the following kind of argument warrants explaining Ashwin's action by reference to a particular consideration in favour of it:

(M1) Ashwin intends to cheer Sarah up.

(M2) Ashwin acts on the belief that Sarah likes lilies.

Therefore,

(M3) Sarah's liking lilies is among Ashwin's reasons for buying Sarah lilies.

Not all Mentalists think that the contents of the *beliefs* on which Ashwin acts determine his reasons for acting: a Mentalist is just someone who thinks that the content of *some* state of mind determines which reasons Ashwin acts for. So the Mentalist might claim that Ashwin must *know* that Sarah likes lilies in order for this to be a reason for his action. Or the Mentalist might insist that Ashwin must believe that Sarah's liking lilies is a consideration in favour of buying them for her, rather than just registering the bare fact that she likes them. Another kind of Mentalist might hold that Sarah's liking lilies is a reason for Ashwin's action only if Ashwin acts in part because he is *concerned* that this fact should explain why he acts.⁴ At the end of this part (§I), I'll briefly consider which kinds of mental states could determine what an agent's reasons for acting are. To simplify things, I'll assume for now that the Mentalist thinks that an agent's beliefs determine his reasons for acting.

What exactly is the Mentalist claiming? She is *not* claiming that the fact that Ashwin has the belief mentioned in (M2) justifies his action. Rather, she is claiming that the content of this belief can be used to justify his action. This is, I think, uncontroversial. The Mentalist's controversial claim is that citing this particular content in justifying Ashwin's action can explain why it happens *only because* Ashwin acts on a belief with this content. This is a theory about what it is to act for particular reasons. The content believed justifies the action, but it is the fact that Ashwin believes it that makes it explanatory to cite this content.

Is the Mentalist's claim really controversial? In ordinary life, it would seem very strange to explain Ashwin's action in terms of things we knew he didn't believe. So most philosophers would probably agree that a way of justifying Ashwin's action is

only relevant to explaining his action if he actually believes the claims by which his action is justified. But this isn't sufficient to show that the Mentalist is right. At most, it shows that if we can give an explanatory justification of Ashwin's action by reference to some content, then Ashwin must have a belief with that content. The Mentalist's claim is stronger: she claims that it is *because* Ashwin believes this content that it is explanatory to cite it in justifying his action. More generally, the Mentalist claims that it is Ashwin's state of mind that determines why it is rational (or irrational) for him to act as he does.

The Teleologist offers a radical alternative to the Mentalist's position. The Teleologist denies that an agent's state of mind determines which reasons explain his action. Instead, she thinks that an agent's reasons for acting are determined by the teleological mechanisms responsible for his actions. To put the view in its simplest form, a consideration is among Ashwin's reasons for buying Sarah lilies if the mechanism producing this action is designed so as to achieve Ashwin's purpose only on condition that Sarah likes lilies. That is, something counts as a reason for an action in virtue of being a normal condition for the proper functioning of the mechanism responsible for that action. How does this view work? The key notions are 'proper functioning' and 'normal condition'. Different Teleologists give different accounts of these notions. One kind of Teleologist defines the proper function of an action etiologically, in terms of facts about how actions have historically served purposes and in virtue of which they have been reproduced. A normal condition for the proper functioning of an action is a condition that must be mentioned in explaining why the historical antecedents of that action were successful. This is (a much simplified version of) Ruth Millikan's approach. On her view, Ashwin's buying lilies for Sarah has the function of cheering her up because earlier instances of this type of action have succeeded in cheering people up, and because these historical successes explain why Ashwin now buys lilies. Sarah's liking lilies is a normal condition for the success of Ashwin's action because the successes of those earlier lily-buying exploits in virtue of which Ashwin is now acting are explained in part by the fact that the recipients liked lilies. And Sarah's liking lilies is a reason for Ashwin's action because it is a normal condition for it to function properly.⁵ Another kind of Teleologist defines the reasons for an action in terms of what that type of action is theoretically suited to bringing about, rather than

in terms of what it has historically brought about. This is Rowland Stout's approach. On his view, Sarah's liking lilies is Ashwin's reason for buying them if his action is theoretically suited to cheering Sarah up given that Sarah likes lilies. More generally, an agent's reasons for acting are those conditions that are required for the correct theoretical description of his action to describe a process that is suited to achieving his purpose in acting.⁶

Despite their differences, I group these two kinds of Teleologist together because they both hold that facts about the function of an action are what makes that action appropriate, and they both regard the function of an action as something that can be grasped without first knowing anything specific about the contents of the agent's mind.⁷ The Teleologist holds that the following kind of argument determines which considerations in favour of Ashwin's action are his reasons for acting and therefore explain why he did it:

(T1) Ashwin intends to cheer Sarah up.

(T2) Sarah's liking lilies is a condition for Ashwin's action to fulfil his intention normally.

Therefore:

(T3) Sarah's liking lilies is among Ashwin's reasons for buying Sarah lilies.

So the Mentalist and the Teleologist might agree that Ashwin's action is explained by the consideration that Sarah likes lilies, but they will disagree about why this consideration is relevant to explaining his action. Is it *belief* or *function* that determines the relevance of a consideration to explaining an action? For the Mentalist, a consideration is explanatory in virtue of something Ashwin believes; for the Teleologist, the consideration is explanatory in virtue of being a condition for the teleologically normal functioning of Ashwin's action. In short, the Mentalist thinks that facts about Ashwin's state of mind determine what his reasons for acting are, and the Teleologist thinks that it is facts about the functions of Ashwin's actions that determine which reasons he acts for.

2. The difference between the Teleologist and the Mentalist can appear to be very slight. We can think of them as proposing different ways of identifying which

justifications count as explanations of why an action happened. This suggests that we could decide which of them is right by comparing what is actually explanatory in different situations with what each theorist implies should be explanatory. But this won't work, because although they disagree about why certain statements are explanatory, they need not disagree about which statements are explanatory. To show this, I will now explain how the Mentalist and the Teleologist could come to agree in every situation about which considerations explain why an agent acts.

For the Teleologist, facts about Ashwin's reasons for acting don't rest upon facts about what he thinks. This allows the Teleologist to analyse facts about what Ashwin believes *in terms of* facts about the reasons for his action. Different Teleologists do this in different ways. Millikan, for example, identifies the content of a belief with the conditions under which actions caused by that belief should achieve their purposes. To simplify, Millikan can say that if Ashwin has a belief that causes him to buy Sarah lilies when he intends to cheer her up, and if it is a normal condition for the success of his action that Sarah likes lilies, then his belief is the belief that Sarah likes lilies.⁸ In this way, the Teleologist can explain what Ashwin believes in terms of what he should do:

(T1) Ashwin intends to cheer Sarah up.

(T2) Sarah's liking lilies is a condition for Ashwin's action to fulfil his intention normally.

Therefore:

(T3) Sarah's liking lilies is among Ashwin's reasons for buying Sarah lilies.

Therefore,

(T4) The belief on which Ashwin acts is the belief that Sarah likes lilies.

I think Millikan would endorse the inference to (T4) as an explanation of what it is for Ashwin to believe that Sarah likes lilies because she regards the contents of beliefs as determined by facts about what agents should do.⁹ Although Stout takes a different approach, he also analyses facts about beliefs in terms of facts about what agents should do.¹⁰ Note that a Teleologist who explains belief in this way isn't just claiming that the inference to (T4) is valid. Rather, the Teleologist is claiming that

the inference to (T4) explains what it is for a belief to have a particular content. For her, facts like (T1) to (T3) are canonical grounds for ascriptions of belief.

Although Teleologists like Millikan or Stout endorse teleological accounts of belief along these lines, a Teleologist doesn't have to do this. As I define her, the Teleologist makes a claim about action, not belief: she claims that which reasons explain an action is determined by some aspect of the teleological function of that action. It is not obviously incompatible with this to deny that which beliefs a person has can be explained in terms of the teleological functions of the actions they result in. So my point is only that the Teleologist's account of action makes it possible for her to give a theory of belief in terms of action. This possibility is not available to the Mentalist, however, because the Mentalist thinks that which reasons an agent acts for – and therefore which actions he performs – is determined by what he believes. The Mentalist's notion of action thus presupposes facts about what an agent believes, so it would be circular for her to give an account of belief in terms of action. She may accept as valid and even informative the kind of reasoning in the Teleologist's inference (T1)-(T4), but she can't regard this pattern of reasoning as explaining what belief is. If an agent's beliefs are to determine which reasons he acts for, then what it is for a belief to have a particular content cannot be defined in terms of the actions it should result in.¹¹

If the Teleologist does define belief in terms of teleologically normal conditions for successful action, she will hold that if a consideration is among an agent's reasons for acting, then it follows that he acts on a belief concerning that consideration. The Teleologist only denies that this consideration is the agent's reason for acting *in virtue of* his acting on the belief. She will therefore regard the Mentalist as invariably right about which things are an agent's reasons for acting, but wrong about what makes it the case that certain considerations in favour of an action are the reasons that explain why it happens.

For her part, the Mentalist can agree with the Teleologist that a consideration is among an agent's reasons for acting only if it is a condition of his action fulfilling its purpose normally. This is because the Mentalist can define normal conditions for an

action to succeed in terms of the contents of the agent's beliefs. For example, the Mentalist can reason as follows:

(M1) Ashwin intends to cheer Sarah up.

(M2) Ashwin acts on the belief that Sarah likes lilies.

Therefore,

(M3) Sarah's liking lilies is among Ashwin's reasons for buying Sarah lilies.

Therefore,

(M4) Sarah's liking lilies is a condition for Ashwin's action to fulfil his intention normally.

This shows that the difference between Teleologist and Mentalist isn't necessarily discernable in terms of which reasons each theorist will regard as explaining why an action happened. Instead, the key difference between the Teleologist and Mentalist is a difference about *the order of explanation*. The Mentalist holds that Ashwin's state of mind determines what he should do, whereas the Teleologist thinks that Ashwin's state of mind is determined by the facts about what he should do.

3. How is this conflict between the Teleologist and the Mentalist relevant to a distinction between two kinds of purposive action? At the start of this discussion, I said that the Teleologist and the Mentalist each describe a different kind of rationalising explanation for purposive actions. However, as I have characterised their disagreement so far, it might seem a step removed from the question of what kind of explanation is appropriate for explaining purposive action. The Teleologist and Mentalist appear to be assuming a common model of explanation, namely the model of ordinary practical reasoning. The difference between their positions seems to concern not what kind of reasoning explains an action, but only what sort of fact determines whether a particular instance of this kind of reasoning can explain why an action happens. However, the difference between Teleologist and Mentalist does in fact have consequences for the kinds of reasoning that explain why agents act. This is because their debate concerns the nature of norms for explaining purposive action. The Teleologist and the Mentalist agree that explanations of action invoke norms. For example, an explanation of why Ashwin buys Sarah lilies must reveal why buying Sarah lilies is something Ashwin should do given that he intends to

cheer her up. But the Teleologist and Mentalist disagree about the kind of norm involved: for the Teleologist, the norm is teleological, whereas, for the Mentalist, the norm is mental. This means that the Teleologist and the Mentalist understand the claim that Ashwin *should* buy Sarah lilies differently.

The Teleologist explains why there is something that Ashwin should do in terms of facts about the history or theoretical suitability of types of action. These are facts that you can comprehend without having any insight into what Ashwin himself takes to be reasons for acting, and these facts don't presuppose that Ashwin thinks about reasons at all. The only part Ashwin himself plays in making it the case that his action is subject to particular norms is purposing to cheer Sarah up by buying her lilies. The requirements for his action to succeed normally, such as the requirement that Sarah likes lilies, are imposed by facts about the history or theoretical suitability of his action. He must be somehow sensitive to these requirements, but they are imposed on his actions because of their teleology and not because he is aware of them as norms governing his action. So, on the Teleologist's account, norms are of no concern to the agent; or if the agent does happen to be concerned about which norms explain his action, this has no bearing on the question of whether these norms govern his action.

If an agent is not at all concerned about his reasons for acting, the Teleologist's view is probably a natural one. But the Mentalist objects that the Teleologist ignores the fact that purposive actions are done by agents with their own perspectives on why they should act. Everyone agrees, for the purposes of this debate at least, that we are often aware of reasons for our actions as reasons. The Mentalist asks what the significance of our being aware of reasons is. Her answer is that an agent's awareness of reasons partly determines which action he performs. Of course, an agent's awareness of particular reasons doesn't determine that he will in fact act in accordance with those reasons, but it does determine that these reasons are the right standards by which to judge his action. The insight a Mentalist's account of action should aim to capture is nicely put by Davidson:

'norms ... enter in a special and additional way in the study of mental phenomena. Whatever is studied, the norms of the observer will be

involved. But when what is studied is the mental, then the norms of the thing observed also enter.' (1990, 25)

I take it that 'the norms of the thing observed' are those norms which the thing observed is aware of as norms applying to him, or is concerned to act in accordance with. According to Davidson, then, giving a mental explanation requires us to appreciate the agent's own perspective on the reasons for his action. Particular norms apply to an action in virtue of the agent's appreciation of the relevance of these norms to his action. So his actions need to be explained in terms of his mental states on the model indicated by the Mentalist.

This gives us the link between the nature of particular explanations of action and the more abstract debate between Teleologist and Mentalist about what makes an explanation explanatory. In explaining an action, a Teleologist and a Mentalist might use the same words. But they would be using these words to invoke quite different kinds of normative standard, and with correspondingly different ideas about the point of explaining why an agent acts. The Teleologist's explanation explains why an action is the sort of thing that normally fulfils the agent's purpose in acting; the Mentalist's explanation explains why an action is rational from the agent's point of view. So the Teleologist and the Mentalist give us two different models of explanation for purposive actions, and therefore two notions of action.

4. The Teleologist and the Mentalist are usually taken to be offering conflicting accounts of a single phenomenon. However, if we accept that there are different kinds of purposive action, we can regard the Teleologist and Mentalist as describing different schemes of explanation appropriate to different kinds of purposive action. Or, to turn this around, the existence of an apparently irresolvable dispute between the Teleologist and the Mentalist suggests that they might be talking about different things, and this is circumstantial evidence for the existence of more than one kind of purposive action. So perhaps the accounts of rationalising explanation given by the Teleologist and the Mentalist should be regarded as accounts of the types of explanation appropriate to two kinds of purposive action.¹²

Suppose that the Teleologist and Mentalist could be interpreted as identifying two models of rationalising explanation and thereby providing a theoretically viable distinction between two kinds of purposive action. Would this distinction relate to the intuitive distinction between actions done for reasons and information-based but unreflective actions that I mentioned at the start? Take information-based but unreflective actions first. I suggested that information-based actions are reasonable and intelligent ways of achieving goals that don't depend on the agent's ability to understand his reasons for acting. This fits with the Teleologist's notion of rationalising explanation, which doesn't require that the agent have thoughts concerning why he should act. (Although some Teleologists have claimed that agents are necessarily aware of their reasons for acting, this isn't integral to the Teleologist's position.) So the Teleologist's position can naturally and plausibly be regarded as an account of a kind of purposive action that does not require the agent to be aware of his reasons for acting.

Is there also a connection between the Mentalist's account of explanation and the intuitive notion of an action done for a reason? I think there is, but the connection is not obvious. An action is done for reasons when the agent is aware of his reasons for acting *as reasons* and acts in part because of this awareness. Minimally, a reason is the kind of thing that can be used to give a practical justification for acting. So being aware of something as a reason for acting involves having some understanding of how it can justify the action.¹³ Now it is not clear why the Mentalist's account of action explanation applies only to actions where the agent has this reflective awareness of reasons. As I defined her, the Mentalist is anyone who holds that the contents of an agent's states of mind determine which considerations are his reasons for acting. In characterising the Mentalist's position, I made no assumptions about which states of mind determine what an agent's reasons for acting are. Clearly, not every state of mind involves having a reflective understanding of its content as a potential reason for action. So it seems that the Mentalist might not accept that an agent is necessarily aware of his reasons for acting *as reasons*. For couldn't the Mentalist claim that a consideration is an agent's reason for acting in virtue of the fact that an agent acts on a belief with that content, but independently of whether the agent is aware of the content of his belief as a reason for his action? If she can, her account of explanation won't fit the

intuitive notion of an action done for reasons, because it will also apply to actions done without awareness of why the action is rational.

To show that the Mentalist's account of action explanation does fit the intuitive notion of an action done for reasons, I need to argue that whatever state of mind the Mentalist thinks determines an agent's reasons for acting, having that state of mind must imply that the agent is aware of its content as a reason for action. I'll first argue that merely acting on a belief is not sufficient for the content of that belief to be a reason for the action. The reason for this is straightforward: agents sometimes act on beliefs the contents of which are not among their reasons for acting. Instances of this are exceptional because agents rarely act on a belief unless they also appreciate how that belief bears on their objectives, but here is one. Imagine that Richard shows special favour towards one of his subordinates in order to get her to think well of him, and that he does this because he believes she is very wealthy. If Richard thought that wealthy people tend to respond well to favouritism, his subordinate's wealth might be a reason for his favouring her. However, Richard believes that favouritism won't get his subordinate to think well of him, because he thinks wealthy people are beneath being impressed by such things. But he can't help himself: the impression of his subordinate's vast wealth combined with a desire for her good opinion invariably impel him to acts of favouritism that, by his own lights, thwart his purpose. So although Richard acts because he believes that his subordinate is wealthy, this is not a reason for his action.¹⁴

This example indicates that merely acting on a belief isn't a sufficient condition for the content of that belief to be a reason for the action.¹⁵ However, the example doesn't quite explain why a Mentalist can't claim that an agent's beliefs determine his reasons for acting without the agent necessarily being aware of their contents as reasons for acting. It is often suggested that cases like Richard's are deviant because the belief does not result in action 'in the right way'. Superficially, then, it seems a Mentalist could claim that an agent's beliefs determine which reasons he acts for, but only when his beliefs result in action in the right way. But what does 'in the right way' mean? What is the right way for a belief to result in action? Intuitively, the problem with cases like Richard's is that an agent's actions are not always *rational* consequences of his beliefs. It is only plausible that the contents of

an agent's beliefs are reasons for his action when his actions happen because they are rational consequences of the beliefs he acts on. So an account of what it is for a belief to result in action 'in the right way' is, in effect, an account of what makes it the case that the content of that belief is among the agent's reasons for acting. But now we can see that the idea that the contents of an agent's beliefs determine his reasons for acting can be left out of the picture altogether. Someone who says that the contents of an agent's beliefs determine his reasons for acting when he acts on them 'in the right way' could put her position more transparently. Instead of saying that an agent's beliefs determine which reasons he acts for, it would be more accurate for her to say that some further fact makes it the case that the contents of an agent's beliefs coincide with his reasons for acting. The 'further fact' is whatever condition is required for an action to result from a belief in the right way, and it is this fact that makes it the case that a consideration is among an agent's reasons for acting. This position is purely schematic, however. The Teleologist can accept this view, providing she gives a teleological account of what it is for an action to result from beliefs in the right way. The Mentalist can also accept this view, but she will surely have to explain 'the right way' for beliefs to result from action in terms of the agent's awareness of the contents of his beliefs as reasons for action. So we are right back with the question I started with: What determines which considerations are among an agent's reasons for acting? I conclude that the Mentalist can't claim that an agent's awareness of a consideration as a bare matter of fact makes it the case that this consideration is among his reasons for acting. Either this claim is false, or else it fails to explain what makes it the case that an agent acts for certain reasons. Either way, an agent's being aware of something as a bare fact doesn't make it a reason for his action.

The same conclusion can be reached from reflection on reasons. We can explain why an agent should or shouldn't act in terms of his reasons for acting. Suppose an agent is aware of some fact that we regard as a consideration against his action, although the agent himself doesn't see how the fact might justify acting or not acting. The agent acts in part because he is aware of this fact. In this situation, we can't criticise the agent's action on the grounds that he was aware of a fact *we* regard as counting against it. At most, we can criticise him for not seeing the import of this fact. By the same token, then, the fact that an agent is aware of a fact *we*

regard as favouring his action can't warrant us in counting that fact among his reasons for acting unless the agent also regards this fact as favouring his action. We may be entitled to justify an agent's action by reference to some fact even if he doesn't see how it justifies his action (as the Teleologist claims), but we can't be entitled to do so *in virtue of* his being aware of it as a bare fact. In other words, if an agent isn't aware of a consideration as relevant to a practical justification of an action, then either this consideration is not a reason for that action, or, if the consideration is a reason for his action, this is not because the agent is aware of it. So unless we are aware of something as a reason for acting – unless we at least see how it could be used to justify acting – our awareness of it can't make it one of our reasons for acting. Therefore, whatever kind of mental state the Mentalist regards as the reason-determining kind, acting on states of this kind must involve having some understanding of how your reasons for acting could justify your action.¹⁶

This is why I think that we can use the Mentalist's model of action explanation to make precise the intuitive notion of an action done for reasons. Intuitively, an action is done for reasons if it is done in part because the agent is aware of reasons for doing it. The Mentalist holds that something is a reason for an action in virtue of its being the content of a mental state on which the agent acts. This theory captures this intuition, for if the content of a mental state is to determine which reasons an agent acts for, then having that mental state must involve awareness of the reason as a reason for the action. So we can define an action done for a reason as the kind of action that is characterised by the Mentalist's account of reason explanation.

To sum up this first part, I have argued that the Teleologist's account of action explanation can be regarded as an account of the kind of explanation appropriate to information-based but unreflective actions, and that the Mentalist's model of explanation applies to actions done for reasons. This shows that the intuitive distinction between these two kinds of purposive action can be made theoretically viable. It doesn't show that there are two kinds of purposive action, however, because it may be the case either that every purposive action can be explained using the Teleologist's model of explanation, or that all purposive actions can be explained on the Mentalist's model. This is what the Teleologist and Mentalist claim. Neither need deny that there are competing theories, but each claims that her

theory is the *right* one for every purposive action. My aim in the next part is therefore to argue that no single scheme of explanation applies to all instances of purposive action.

§II Is the distinction empirically motivated?

Assuming that the Teleologist and the Mentalist provide a theoretical basis for the distinction between information-based but unreflective actions and actions done for reasons, is there any empirical motivation for this distinction? The Mentalist might object that unless an agent can be aware of his reasons for acting as reasons, his actions are not genuinely purposive. They are mere response behaviour; and if they appear to be purposive, that is only because of their complexity. So, in other words, the objection is either that the Teleologist's notion of action isn't genuinely purposive, or that we don't need the notion of purposiveness that the Teleologist aims to capture.¹⁷

This objection shows how counterintuitive the Mentalist's position is. It seems obvious that we sometimes do complex and novel things purposively but without any clear idea about why we should do them. 'I don't quite know why I did that,' we might say, 'but it worked.' Reasons don't concern us: we just get on with the doing. We are sometimes even surprised when we realise why we are doing something, or so it seems. According to the Mentalist, however, either an action is not purposive or we know why we do it.¹⁸ Of course, it is no reply to the Mentalist just to say that this isn't how things seem.

To reply to the Mentalist's objection, we need to ask what is required for an action to be purposive. A standard answer is that genuinely purposive behaviour occurs when an agent has a goal, and when the agent's having this particular goal might result in any one of a range of actions, and when the agent's having the goal normally results in whichever action is the most appropriate way of achieving the purpose (cf. Taylor 1964, 9; Bennett 1976, §12). If this is all that is required for an action to be genuinely purposive, it doesn't seem very likely that purposive actions require the agent to be capable of being aware of the reasons for his action. Why should this kind of awareness be necessary for an agent to be able to perform any

one of a range of actions depending on which is the most appropriate way of achieving his purpose?

This becomes clearer when we look at how we could test whether a simple kind of behaviour is purposive. Take a rat that has been trained to press a lever in order to dispense sugar solution. When the rat presses a lever, some sugar solution drips into a bowl. After pressing the lever repeatedly, the rat can approach the bowl and drink the sugar solution. Is the rat's lever-pressing a purposive action or a response? If the lever-pressing action is a response, then it can be explained in associationist terms along the following lines:

The rat learns to associate pressing the lever with obtaining sugar solution. The rat then presses the lever because it has a preference for drinking sugar solution.

If, on the other hand, the rat's action is a purposive action, then it needs a rationalising explanation along the Teleologist's lines:

The function of the rat's lever-pressing is obtaining sugar solution. The rat intends to get sugar solution, so it should press the lever.

The Mentalist's way of explaining purposive action can't be applied to the rat. This is not because rats don't have beliefs (maybe they do). It is because reason explanation, as the Mentalist understands it, requires that the agent is aware of his reasons for acting as reasons, and it doesn't seem very likely that rats can think about their reasons for acting. So if the rat's action can be shown to be purposive, the Mentalist should concede that the Teleologist's notion of action is purposive.

How can we tell whether the rat's lever-pressing is a purposive action or merely a response? Anthony Dickinson and Bernard Balleine have invented a test, which shows that the rat's behaviour is not controlled by associations. Here is a simplified description of it (see Dickinson and Balleine, 1993; Dickinson 1994). In the first stage of the test, they put a rat in a room with a lever, which, unknown to the rat, dispenses sugar solution. Initially, the rat presses the lever occasionally, as if out of curiosity. But once it discovers the sugar solution, which it likes to drink, it quickly learns to associate pressing a lever with obtaining sugar solution. It presses the

lever frequently to get as much sugar solution as it can. The experimenters then make the rat ill, which causes it to become averse to the sugar solution. Once rats have been given this kind of treatment, they will tend to avoid going near sugar solution. But this particular rat, after being made averse to the sugar solution, is kept well away from sugar solution. Once it has recovered from its illness, the rat is given the opportunity of pressing the lever again. This is the second stage of the test. What does the rat do? A basic principle of associationist explanation is that actions associated with positive outcomes are performed relatively often, whereas actions associated with negative outcomes are performed relatively rarely. The rat associates pressing the lever with obtaining sugar solution, but now it is averse to sugar solution. So an associationist explanation of the rat's lever-pressing behaviour would require it not to press the lever very much. However, this isn't what happens. Instead, the rat initially presses the lever just as vigorously as another rat that actually liked sugar solution would. So the rat appears to do exactly the opposite of what an associationist explanation says it should do. Once some sugar solution has been delivered, the rat immediately manifests its aversion by not going near it. Furthermore, once the first lot of sugar solution is delivered and the rat has manifested its aversion, it then desists from pressing the lever. This is also problematic for an associationist explanation of the rat's behaviour. For since the rat was averse to sugar solution from the time of its illness, encountering the sugar solution again shouldn't make a difference to its lever-pressing behaviour.

What is the rat doing? Intuitively, we might think that the rat initially presses the lever after being ill because it doesn't *know* it is averse to the sugar solution and still wants it. Then, once it has seen the sugar solution, it realises that it is averse to it and so stops pressing the lever. Thus the rat acts with the purpose of getting sugar solution *in spite of* being averse to it. This shows that the rat's behaviour can easily be accommodated by a rationalising explanation, for it is perfectly coherent to want something to which one is averse. However, this way of thinking has no place in associationist explanation, because the point of this kind of explanation is to explain behaviour in terms of exposure to contingencies without reference to more sophisticated kinds of experience or knowledge. It is not being aware of an aversion but the aversion itself that is supposed to inhibit behaviour.

The rat's behaviour seems impossible to explain in associationist terms. The basic problem is that, at the start of the second stage of the test, the rat presses the lever as vigorously as it would if it assigned positive value to the associated outcome, whereas it is actually averse to it. In order to fit the rat's behaviour into an associationist explanation, we might postulate that the rat doesn't become averse to the sugar solution until after re-encountering it; but this would be anomalous and there's no independent evidence for it. Alternatively, we might conjecture that the rat's aversion to sugar solution somehow fails to have its normal inhibitory effect on the rat's actions until the rat re-encounters sugar solution after being ill. However, this conjecture doesn't explain why the rat presses the lever so frequently – as frequently as a rat that actually liked sugar solution would. To explain this in associationist terms, we have to assume that the rat assigns positive value to the outcome it associates with lever-pressing. But we can't do this, because the rat associates lever-pressing with sugar solution to which it is now averse. There is certainly no simple way of fitting the rat's behaviour into an associationist explanation. (Of course the actual tests use sophisticated controls that I have ignored.) It is an empirical question exactly how associationist mechanisms work in rats, so it is possible that someone might find a way to explain the rat's lever-pressing as a response behaviour.¹⁹ But since there is no reason to assume *a priori* that this will become possible, inference to the best currently available explanation requires us to regard the rat's lever-pressing as a genuinely purposive action. So my reply to the Mentalist's objection is that we need a category of information-based but unreflective actions because animals without any capacity to think about reasons as reasons can nevertheless act purposively.

The force of this reply needn't rest entirely on the rat. Assume for the sake of argument that we *can* explain the rat's lever-pressing in associationist terms, which means that we shouldn't regard it as a purposive action. The methodology of the rat experiment is still significant for the issue of whether purposive actions can be done without awareness of reasons. The basic idea of the experiment is this: if an action would continue to be performed although its outcome were devalued, then that action cannot be controlled by associations. Dickinson and Balleine put this point by saying, '[purposive] actions and responses can be dissociated under goal revaluation brought about by ... aversion training' (1993, 291). One characteristic

by which genuinely purposive behaviour can be recognised, then, is that it is behaviour that can occur because it is a way of achieving something the agent desires even if he is averse to it.²⁰ Since this test for purposive behaviour does not appear to presuppose much cognitive sophistication, its adequacy strongly suggests that engaging in purposive action does not always require that an agent understands which reasons he is acting for. Or, to put the point another way, the fact that associationist explanation stops at the point where an agent's goals come apart from his aversions indicates that there is a significant gap between the most advanced response behaviour and the simplest kinds of action that the Mentalist can explain. So even if the rat's lever-pressing isn't purposive, we still have good reason to think that the Mentalist's model of explanation cannot apply to all purposive actions.

Now I face an objection from the Teleologist, who might say that the Mentalist's notion of explanation makes purposive action into something too intellectual. The Teleologist thinks that all purposive actions are information-based but unreflective. So it is wrong to claim, as the Mentalist does, that some purposive actions are done in part because the agent finds them reasonable.

To reply to the Teleologist and to show that being aware of the reasons for an action can affect its character, it suffices to cite a purposive action that the Teleologist's model of explanation cannot accommodate. Just as Dickinson and Balleine show that the rat's lever-pressing is genuinely purposive by showing that it can occur *in spite of* the rat's associations, I need an action that is purposive but done *in spite of* the teleologically normal conditions for its success. What kind of action could this be?

To answer this question, we need to consider the Teleologist's position concerning awareness of reasons. Since her account of purposive action involves no reference to the agent's state of mind, the Teleologist must deny either that we are ever aware of reasons for our actions, or that we ever act because we are aware of our reasons for acting. Although it isn't easy to say what awareness of reasons is, it's hard to see why a philosophical account of action should deny that it exists. On the other hand, if the Teleologist accepts that we are aware of reasons for acting in a way that rats aren't, she has to deny that this awareness makes a difference to the nature of

our actions. For the Teleologist, awareness might still be useful in the long run as an optional extra that enables us to monitor which reasons are producing our actions, but it doesn't immediately affect what we do. This implies that the relationship between action and awareness of reasons is, at best, one of pre-established harmony. At best, we act in accordance with reasons we are aware of. The primary role for awareness of reasons is not active, but passive: it's function is not to set standards for our actions, but to detect whatever standards govern our actions.

This view is counterintuitive. As an analogy, consider an agent's objectives rather than his reasons. Suppose John shoots a rabbit. There are lots of standards by which we might evaluate the success of his shot. Was the rabbit big enough for the family's dinner? Does it have good fur? Was it from an overpopulated warren? However, if John was aiming at a particular rabbit he saw, it is natural to think that standards such as these are secondary for understanding his action. The basic kind of success is hitting the rabbit aimed at. Any other standard is derivative, and concerns not whether John *hit* the right rabbit, but whether he was *aiming* at the right rabbit. An explanation of John's action has to hit the mark he aims at, even if his shot misses. This seems no less true of an agent's reasons for acting. Agents are sometimes concerned that their actions should be justifiable by reference to certain considerations: they have specific ambitions in the arena of reason. When this is the case, it is plausible that there is no more basic standard by which to judge the rationality of an action than the standard the agent himself was aiming at. This is what the Teleologist denies. For her, the right standard by which to explain an agent's actions may coincide with the one the agent was aiming at, but it is not the right standard in virtue of the agent's aiming at it.

The Teleologist's view about awareness of reasons for acting means that her account of action explanation won't be able to accommodate an action where the agent is concerned about which reasons should explain his action, and his concern is in conflict with the teleological function of his action. An instance of this is an irrational action where an agent acts in a way that he knows is not appropriate for achieving his purpose. I think this is a common form of irrationality. For example, imagine that Tom is driving Ayesha in his car and intends to impress her. He attempts to do this by a reckless overtaking manoeuvre, even though he knows that

bad driving won't impress Ayesha. His action is akratic because it is done for a purpose to which he knows it isn't suited.

We must regard Tom's action as genuinely purposive because otherwise it is unclear why what he does is irrational as opposed to merely non-rational. That is, we must distinguish Tom's calculated but self-defeating action from the case of someone whose desire to impress Ayesha causes him to become so frenzied that blind impulses seize him and he overtakes with no purpose at all. On the other hand, in order to see Tom's action as genuinely irrational, we must also distinguish it from the case of someone who overtakes merely because he falsely believes that this is a way to impress Ayesha: this is misguided but not irrational. The problem for the Teleologist is that her theory doesn't allow for these distinctions, so she can't regard Tom's action as purposive but irrational. According to the Teleologist, an agent's reasons for acting are the normal conditions for the success of his action. If there are no normal conditions for the success of an action, then the Teleologist's model of explanation is inapplicable and she can't regard that action as purposive. Because an agent's reasons are normal conditions for the *success* of an action, they will always favour his action. In order to describe an irrational action, however, we need to allow that the reasons which should explain why an agent acts don't always favour his action. For example, Tom's action would not be irrational unless Ayesha's being unimpressed by his overtaking were among the reasons which should explain this action. Since Ayesha's being unimpressed couldn't be a normal condition for the success of his action, the Teleologist's view implies that it isn't relevant to explaining his action. The Teleologist's notion of explanation therefore fails to apply to Tom's overtaking.

The Mentalist, on the other hand, has no special problem with Tom's action, because her theory doesn't involve identifying reasons with teleologically normal conditions for successful action. For the Mentalist, Tom's action is irrational precisely because he fails to act in accordance with considerations that, by his own lights, explain what he should do. (This isn't to say that irrationality isn't philosophically problematic; only that it poses a special problem for the Teleologist.) So the Mentalist's notion of action can accommodate a kind of purposive action that the Teleologist's can't.

This example shows that people sometimes do things for reasons which don't coincide with any of the teleologically normal conditions for the success of their actions, and therefore that there are some purposive actions to which the Teleologist's model of explanation doesn't apply. Although the example uses an irrational action to show that an agent's reasons for acting don't necessarily coincide with normal conditions, it can't be concluded that the Teleologist's model of explanation fails to apply only to irrational actions. This would be like assuming that the rat's lever-pressing behaviour is controlled by associations except when the rat is averse to the outcome of lever-pressing. The fact that the reasons explaining why an agent acts don't coincide with normal conditions in the case of irrational actions shows that his actions can be controlled by non-teleological reasons even when it happens that his reasons for acting do coincide with those that the Teleologist would use to explain his action.

Conclusion

To sum up, I have suggested theoretical and empirical grounds for distinguishing two kinds of purposive action:

- **information-based but unreflective actions**
behaviour that occurs because the agent has a purpose, but where the agent does not act because he is aware of his reasons for acting
- **actions done for reasons**
behaviour informed by the agent's awareness of his reasons for acting

My argument was that some, but not all, purposive actions are done because of the agent's awareness of reasons for his action, and that no single scheme of explanation can cover both types of purposive action. Actions done for reasons occur in part because the agent is aware of his reasons for acting. These actions must be explained from the point of view of the agent himself, in terms of reasons of which the agent is aware *as reasons*. In the case of information-based but unreflective actions, on the other hand, there is no such awareness to appeal to, so explanations of these actions depend only on reasons that apply independently of the agent's

conscious states. The Mentalist and Teleologist's accounts of reasons for action give us models for each type of explanation. They are each right about an important category of action, but neither model of explanation could apply to every purposive action.²¹

Stephen Butterfill
St. Catherine's College
Oxford OX1 3UJ
UK
stephen.butterfill@philosophy.ox.ac.uk

NOTES

- ¹ Of course, not everyone thinks there is a difference in kind between explanations appropriate to behavioural responses and explanations appropriate to purposive actions. Millikan, for example, attempts to locate them both within the broad category of 'teleofunctional explanation', and implies that there is no hard and fast distinction between genuinely purposive behaviour and behaviour with a biological function; cf. e.g. Millikan (1993, 228).
- ² The assumption that we ordinarily explain actions by saying why they should happen is standard but not uncontroversial. It implies that action explanation involves normative considerations: an actual action is explained by reference to an ideal. Some philosophers deny that we ordinarily explain action by appeal to ideal standards. Stich, for example, says, 'the notion of idealized rationality plays *no role at all* in my account [of content ascription] ... we measure others not against an idealized standard but against ourselves. ... this [is] Protagorean parochialism' (1982, 57).
- ³ This question would still be important even if there were only one possible consideration in favour of an action and only one way of justifying it, because we would still need to know whether this consideration explains the agent's action or whether he acts for no reason.
- ⁴ This view is related to one defended by Velleman (1992, 478), who argues that '[w]hat animates practical thought is a concern for acting in accordance with reasons.'
- ⁵ Cf. Millikan (1993, 190): 'explanations of behaviors by reference to reasons for action are normalizing explanations.' I have given a simplified sketch of Millikan's view. In particular, I have described her view as if it required that actions of very narrow types – like buying lilies – each have a specific history in order to have normal conditions. Of course, Millikan's view doesn't require this. On her view, purposive action is produced by a general-purpose belief-desire mechanism. This mechanism provides common causes for different kinds of purposive action. Two actions caused by the same belief will have a corresponding normal condition in common, and actions caused by related beliefs will have correspondingly related normal conditions. In this way, the history of one action can impose normal conditions on other kinds of action. It is important to note that the *contents* of beliefs aren't mentioned in Millikan's account of which normal conditions an action has (i.e. which reasons it is done for). In fact, which content a belief has is defined in terms of normal conditions for the success of actions it results in (see below). As I read Millikan, the postulation of a belief-desire mechanism is just a way of applying her general view about normativity to the specific case of purposive action. The essential point is that a reason for acting is a normal condition for the success of that action. For the full picture, see Millikan (1983, chapters 1 and 2) on proper functions and normal conditions, and Millikan (1993, chapters 7-9) for how she applies these concepts in giving a theory of action-explanation.
- ⁶ See Stout (1996) and (1998), and especially his definition of *circumstantial reason* (1996, 78).
- ⁷ Of course, the most obvious difference between Millikan and Stout probably isn't their different notions of teleology. It is probably that Millikan, unlike Stout, presents herself as upholding the

standard story about beliefs and desires causing action. But one way of putting the point I am trying to argue for in this part is to say that some questions about explanation are not decided by standard claims about the causal structure of action. (Contrast, for example, Child's (1994, 97) claim that by 'seeing that reason explanation is a form of causal explanation ... we understand what sort of explanation it is, and *how* reasons explain actions.')

- ⁸ Strictly speaking, for Ashwin's belief to be the belief that Sarah likes lilies, Sarah's liking lilies must be a normal condition for the success of *all* actions that this belief causes. Cf. Millikan (1993, 71): 'beliefs are ... typed (as beliefs that *p* versus beliefs that *q*) ... in accordance with certain of the conditions that must obtain if they are to fulfill their proper functions (for example, helping to fulfill some-desires-or-other) *in accordance with a Normal explanation*.'

Dennett (1987, 300) aligns his Intentional Stance theory of belief with Millikan's approach: 'it is only relative to such ... 'choices' or evolution-'endorsed' purposes – *raisons d'être* – that we can identify behaviors, actions, perceptions, beliefs, or any of the other categories of folk psychology.' Papineau also advocates a view of belief similar to Millikan's, according to which the truth-condition of a belief is 'that condition which guarantees that actions based on that belief will satisfy the desires it is acting in concert with' (Papineau 1993, §3.6, claim '(C)', my underlining). An important difference between Millikan and Papineau is that Millikan would say 'will Normally' where Papineau says 'will'. For an explanation and defence of this difference, cf. Papineau (1998).

- ⁹ Cf. Millikan (1993, 72-73): 'the truth condition of a belief is a *Normal condition* for fulfillment of [its] proper functions ... That is, the truth condition of a belief is //p. 73// one of the conditions that must obtain if the belief is to fulfill any such functions in accordance with a Normal explanation.' According to Millikan, beliefs feature in two mechanisms, a belief-producing mechanism and a belief-using mechanism. The function of the belief-producing mechanism is to produce belief in accordance with certain 'mapping rules' (i.e. truth-conditions). The belief-using mechanisms perform various functions guided by beliefs in such a way that the belief-using mechanism 'will function properly in the usual way only if the [belief] and the environment match by the relevant mapping rule' (Millikan 1993, 107). So, strictly speaking, which truth-condition a belief has doesn't depend on the particular uses to which that belief is put; rather, it depends on the mapping rules that are Normal conditions for the operation of the belief-producing mechanism. This can make it seem as if which truth-condition a belief has is not defined in terms of its consequences, and thus that Millikan is not necessarily a Teleologist after all. However, although the belief-producing mechanisms are *causally* independent of the belief-using mechanisms, they are *normatively* subordinate to them. What it is Normal for the belief-producing mechanism to do is just whatever the belief-using mechanism needs it to do in order to function Normally. That is, the right mapping-rule for a belief 'is determined by whatever mapping relation is in fact doing the work of successfully guiding the organism through its activities in its world when controlled by the representation' (1993, 108). So there are no facts about which are the right or wrong mapping rules for the belief-producing mechanism independently of the uses to which beliefs are put. The only constraint on which mapping rules the belief-producing mechanism should employ is this: they must be such that, when the mapping rule obtains, the belief-using mechanism is appropriately guided by those beliefs. In other words, the belief-using mechanisms alone set the standard for the correctness of the mapping rules used by the belief-producing mechanism. This is why I regard Millikan as a Teleologist. If this interpretation of her view is wrong, it at least describes a possible Teleologist whose account of action and belief resembles Millikan's.

- ¹⁰ Cf. Stout (1996, 164): 'An agent believes that *P* if and only if a teleological process producing the agent's behaviour is governed by a method of practical justification which in the circumstances works on the assumption that *P* is true'. Stout's account of belief explains what it is for someone to have a belief with a particular content; whereas Millikan's explains, for someone who has a particular belief, what it is for that belief to have a particular content. Unlike Millikan, Stout doesn't assume that we can identify which beliefs are causing which actions independently of knowing their contents.

- ¹¹ It might be objected that the Mentalist *can* give an account of belief in terms of conditions on the success of actions, providing that her account uses only types of action grouped together by their purposes without reference to the reasons for which they are done. On this way of typing

actions, you can do the same thing for different reasons. Since actions are typed without reference to the reasons for which they are done, nothing is presupposed about the agent's beliefs. So the objection is that the Mentalist can give an account of belief in terms of success conditions for these types of action.

This objection fails. If we type actions together regardless of the reasons for which they are done, it is not possible to individuate beliefs in terms of normal conditions for successful action. Suppose, for example, that Ashwin buys lilies for Sarah to cheer her up because he believes she likes them. Ashwin might succeed in cheering Sarah up even if she doesn't like lilies (for example, because she likes being given flowers): this success must be counted as accidental. Equally, Ashwin might fail to cheer Sarah up even if she does like lilies (for example, because she acquires a temporary allergy to them): this failure must also be counted as accidental. These cases of success and failure must be counted as accidental, because otherwise it will not be plausible to say that Ashwin's belief is true if and only if his action succeeds non-accidentally. But these instances of success and failure cannot be counted as accidental if we type actions only by their purposes, ignoring the agent's reasons for doing them. For suppose James buys Sarah lilies because she likes receiving flowers. Acting under the circumstances just mentioned in connection with Ashwin, James' success or failure must be counted as non-accidental. So if we want to give an account of belief-contents in terms of the normal conditions for successful action, we have to use a more fine-grained way of typing actions on which no two actions are of the same type unless both are done for the same reasons. Since the Mentalist explains what it is for an action to be done for a particular reason in terms of its resulting from a belief with a certain content, she can't also explain which content a belief has in terms of which actions that belief results in.

- ¹² Incidentally, I argue elsewhere that the Teleologist and the Mentalist have different notions of *belief* as well as of action. However, the difference between the Teleologist and the Mentalist cannot be explained entirely in terms of differences in their notions of belief, because their different ways of characterising beliefs will presuppose different things about the nature of purposive action. The Teleologist's way of defining belief presupposes the notion of an information-based but unreflective action. If there weren't actions of this kind, her way of explaining belief would be circular. Equally, there is probably no way to explicate the Mentalist's notion of belief without assuming something about the nature of action. Characterising what beliefs are will surely involve reference to the kinds of explanation in which they feature. And since there would be little for beliefs to explain in the absence of any kind of action, a model of explanation for belief must also be a model of explanation for action. So I think it is a mistake to assume that theories about the nature of belief are conceptually prior to, or neutral concerning, theories about what action is.
- ¹³ Note that, for the Mentalist, being aware of one's reasons for acting as reasons needn't involve being aware that one believes these contents. It only involves being aware of the *contents* of one's beliefs as reasons, which doesn't directly imply being aware of beliefs unless reasons *are* beliefs. The Mentalist need not claim that reasons are beliefs: her claim is only that the contents of an agent's beliefs determine which reasons explain his action.
- ¹⁴ This counterexample, and my discussion of it, draws on Velleman (1992), especially §III.
- ¹⁵ It may be true that if an agent has a reason for acting and acts on a belief, then the content of this belief is one of his reasons for acting. (The counterexample involving Richard is consistent with this, because it is quite plausible that he acts for no reason at all.) This is important for a Teleologist who wants to explain what it is for a belief to have a certain content in terms of which reasons explain an agent's actions. It is of no relevance to the Mentalist, however, because she needs to explain what it is for a reason to explain an agent's actions in terms of the contents of the beliefs (or other mental states) on which he acts.
- ¹⁶ This isn't to say that a mental state only provides a norm for action if its content is about that norm. Rather, the point is just that an agent must understand how the contents of the mental states on which he acts count for or against his actions. So the Mentalist can claim that the contents of an agent's beliefs determine which reasons he acts for, providing that she also holds that having beliefs involves being aware of their contents as reasons for action.

- ¹⁷ This seems to be the view Davidson held in 1982, when he wrote 'Rational Animals'. The thesis of 'Rational Animals' is that having 'thought, belief, intention, desire and the like – requires language' (1982b, 322). If Davidson also held that all purposive action is intentional, he would be committed to the view that only language users can act purposively. I think this view is tempting only if one holds that purposive action requires that the agent is aware of reasons for acting as reasons (i.e. that all purposive actions are done for reasons).
- ¹⁸ To deal with cases where we appear to be unaware of our reasons for acting, a Mentalist might say that an agent must at least believe that his way of acting should secure his objective. So we are always aware of a simple reason for acting, namely that our action is appropriate to its objective. However, this doesn't help the Mentalist. The Mentalist holds that something can be a reason for action only in virtue of being the content of a mental state. So if we are aware only that our action is appropriate to its objective, then the Mentalist has to say that the only practical justification we can use to explain why we acted is the generic 'We did it because it works'. In other words, the same reason explains every action done without awareness of any particular reasons for doing it. This is no less counterintuitive than the view that such actions are not purposive.
- ¹⁹ Dickinson and Balleine (1993, §4) cite further experimental support for their position, but they also emphasise that there may be a way of accommodating the rat's behaviour in a more sophisticated model of associationist explanation than any currently available (1993, 291-2).
- ²⁰ Cf. Dickinson and Balleine (1993, 288): 'the content[s] of desires ... have no direct contact with the processes that actually determine whether a particular stimulus or event is desirable, namely whether the stimulus, either directly or via associative links, activates the reward (or punishment) elements in the reinforcement system'. It appears that being able to engage in genuinely purposive action involves an ability to be mistaken about desire. The possibility of this kind of mistake forces us to distinguish desire from its primitive counterpart, affective significance.
- ²¹ An earlier version of this paper was presented at 'Actions', a conference at Manchester University. I am very grateful to the participants, and to Bill Brewer, Constantine Sandis, Jeff Seidman, Rowland Stout and Ros Villa for comments. I am also much indebted to an anonymous referee for suggestions and criticism.

REFERENCES

- Bennett, J. (1976), *Linguistic Behaviour*. Cambridge: Cambridge University Press.
- Brewer, B. (1993), 'Introduction: Action' in N. Eilan R, McCarthy and B. Brewer (eds.), *Spatial Representation*. Oxford: Oxford University Press.
- Brewer, B. (1995), 'Compulsion by Reason', Proceedings of the Aristotelian Society, Supplementary Volume 69, 237-253.
- Child, William (1994), *Causality, Interpretation, and the Mind*. Oxford: Oxford University Press.
- Davidson, D. (1982), 'Rational Animals', Dialectica 36, no. 4.
- Davidson, D. (1990), 'Representation and Interpretation' in Said, M. *et al* (eds.), *Modelling the Mind*. Oxford: Oxford University Press, 1990.
- Dickinson, A. (1985), 'Actions and habits: the development of behavioural autonomy' in Laurence Weiskrantz (ed.), *Animal Intelligence*. Oxford: Clarendon Press.
- Dickinson, A. (1989), 'Expectancy Theory in Animal Conditioning' in Stephen Klein and Robert Mowrer (eds.), *Contemporary Learning Theories*. London: Erlbaum.
- Dickinson, A. (1994), 'Instrumental Conditioning' in N. Mackintosh (ed.), *Animal Learning and Cognition*. London: Academic Press.
- Dickinson, A. & Balleine, B. (1993), 'Actions and Responses: the Dual Psychology of Behaviour' in N. Eilan, , R. McCarthy, and B. Brewer (eds.), *Spatial Representation*. Oxford: Oxford University Press.
- Durkheim, E. (1914/1983), *Pragmatism & Sociology*. Cambridge: Cambridge University Press.
- Gordon, R. (1996), '“Radical” simulationism' in Carruthers and Smith (eds.) *Theories of Theories of Minds*. Cambridge: Cambridge University Press.
- Johnston, M. (1993), 'Objectivity Refigured: Pragmatism Without Verificationism' in John Haldane and Crispin Wright (eds.), *Reality, Representation and Projection*. Oxford: Oxford University Press.
- Millikan, R. (1984), *Language, Thought, and Other Biological Categories*. Cambridge, Mass: MIT Press.
- Millikan, R. (1993), *White Queen Psychology and Other Essays for Alice*. Cambridge, Mass: MIT Press.
- Papineau, D. (1993), *Philosophical Naturalism*, Oxford: Basil Blackwell.
- Papineau, D. (1998), 'Teleosemantics and Indeterminacy', Australasian Journal of Philosophy vol. 76, no. 1, 1-14.
- Perner, J. (forthcoming), 'Dual Control and the Causal Theory of Action: The case of nonintentional action' in N. Eilan and J. Roessler (Eds.). *Agency and Self-awareness*. Oxford: Oxford University Press.
- Stich, S. (1982), 'Dennett on Intentional Systems' in J. Biro and J. Shahan (eds.) *Mind, Brain, and Functionalism*. Norman: University of Oklahoma Press.

- Stout, R. (1996), *Things That Happen Because They Should: a teleological approach to action*. Oxford: Oxford University Press.
- Stout, R. (1998), 'The Evolution of Theoretically Useful Traits', Biology and Philosophy 13, no. 4, 529-540.
- Velleman, D. (1985), 'Practical Reflection', The Philosophical Review, vol. 94, issue 1, 33-61.
- Velleman, D. (1992), 'What Happens When Someone Acts?', Mind, vol. 101, no. 405, 461-481.
- Wright, L. (1976), *Teleological Explanations*. Berkley: University of California Press.