

# EXPLAINING ACTIONS WITH HABITS\*

BILL POLLARD

From time to time we explain what people do by referring to their habits. We explain somebody's putting the kettle on in the morning as done through "force of habit". We explain somebody's missing a turning by saying that she carried straight on "out of habit". And we explain somebody's biting her nails as a manifestation of "a bad habit". These are all examples of what will be referred to here as habit explanations. Roughly speaking, they explain by referring to a pattern of a particular kind of behaviour which is regularly performed in characteristic circumstances, and has become automatic for that agent due to this repetition.

Admittedly, we only use habit explanations in everyday life rather rarely. Standard contexts include drawing attention to some (perhaps objectionable) idiosyncrasy, like nail biting, or when one's habits lead us astray, as in missing a turning off a familiar route. But habits have the potential to explain a vast number of actions. After all, many habits are not bad, but rather helpful. Think of the myriad habits which help in our morning routines of getting up, dressed, breakfasted, and to work. Habits are often not idiosyncratic, but widely shared. Think of habits of etiquette, language use and even reasoning itself. And, as every sports-person and musician knows, most of the time exercising habits does not mislead us, but is of invaluable help, and frees us to turn our attention to other, more important or interesting things. In this light, habits have the potential to explain many of the vast range of actions which we simply take for granted in our everyday lives. Philosophers, have, however, been slow to recognise this potential.

Alvin Goldman may have recognised this when he wrote: "It must be conceded that the analysis of action I have given is not ideally suited for dealing with habit, nor has it been designed with habitual behaviour in mind" (1970, p. 91). Over thirty years on, what we still lack is a clear sense of the theoretical cost of this neglect, which is by no means unique to Goldman's theory. Indeed the whole of contemporary philosophy of action is affected.

---

\* This paper has been accepted for publication in the *American Philosophical Quarterly*, and is due to appear in the January 2006 issue. Copyright © 2005 *American Philosophical Quarterly*. It has benefited from comments by an anonymous *APQ* referee.

This paper conveys a sense of this cost by describing just how habit explanations can and do explain actions. For what is striking about such explanations is that they make no reference to the psychological condition of the agent at all, at least given the common assumption that the agent has some privileged perspective on her own psychological condition. And this represents a challenge to a prevailing intellectualist philosophy of action which assumes such psychological items must always have an explanatory role. It amounts to a refusal to identify the locus of agency with psychology thus conceived. Attention to habit explanation, then, promises not only to enrich our understanding of human agency, but also to de-centre the perennial philosophical difficulties of relating psychology to behaviour.<sup>1</sup>

Quite a bit of conceptual ground-clearing will be necessary. For contemporary action theory presents two main lines of resistance to the very suggestion, platitudinous in everyday life, that habits can explain actions. The first is the tendency to assimilate habitual behaviours with types of behaviour that do not have any important connection to agency. Examples include reflexes, bodily processes, compulsions and phobias. The thought seems to be that when we draw on these explanatory items, we do not implicate the agent, so if habits explain behaviours at all, that only shows that the behaviours being explained are not really actions.

The second line of resistance is found in the commonly held (and intellectualist) view that if a piece of behaviour is worthy of the epithet “action”, it must be “intentional” under some description, and that in turn means that it must be appropriately related to some psychological item. Depending on your view, this could be an act of will, an intention, a reason, a desire, a belief, or some suitably related combination of these states. Since habits, being patterns of behaviour, are not amongst states of this sort, they cannot themselves provide explanations of actions, or at least not of those which can be regarded as “intentional”.

---

<sup>1</sup> It is possible to preserve the link between agency and psychology by expanding the category of the psychological in such a way that it includes the agent’s habits. This would make the present claim rather less striking, but only at the cost of the category of the psychological hiding an even greater diversity of states than it already does. Given that it is proposed in §5 that habit explanation is of a highly distinctive sort, that is reason to resist this assimilation of habits into psychology, at least until it is shown that other “psychological” explanations work in this way.

But neither view is mandatory. It will be argued negatively in §1, that habits are not to be assimilated to reflexes, bodily processes, phobias or compulsions. And it will be argued positively in §2, that habits bear an important, though often neglected, relation to agency, which need not be thought of in psychological terms. §3 proposes that there is still a sense in which exercises of habits can be regarded as “intentional”. In §4 three accounts of the explanatory relation between habits and their exercises are considered and rejected. These are psychological, teleological and causal accounts respectively. In §5 it is proposed that the explanatory relation is instead *constitutive*. This is another way in which habits are distinct from the standard psychological *explanantia*. Finally §6 suggests some implications for the philosophy of action.

## 1. HABITUAL ACTIONS AND OTHER BEHAVIOURS

We begin by making a case for habits being the plausible *explanantia* of actions, things people do, as opposed to being the *explanantia* of mere behaviours, things people suffer.<sup>2</sup> We first see why the assimilation of habits to reflexes and bodily processes is a mistake, and develop the case by clarifying the contrast between habits and addictions, compulsions and phobias.

One of the notable features of habitual behaviours that may lead us to question their status as actions is their automaticity, that is, their lack of any preceding deliberation, choice or effort, or at least none that is apparent. This is a feature they share with other sorts of behaviour, such as reflexes like flinching, or bodily processes like the digestion, and may for that reason seem to belong in the same class of behaviour.

In one way this assimilation is helpful. For nobody is tempted to say that reflex behaviours or movement of the digestion are explicable by psychological items, and this point can be extended to habits. The problem with this assimilation, however, is that when we explain behaviours by means of reflexes or bodily processes, we do not imply that the behaviours in question are actions. On the contrary, such behaviours are not under the

---

<sup>2</sup> Some support for the latter view can be attributed to Ryle (1949, pp. 41-50) and Winch (1958, pp. 57-65). Ryle contrasts “pure habits” with “intelligent capacities”, and argues that it is only the latter that can exorcise the Cartesian ghost. Winch denies that rule-following behaviour is habitual on the grounds that in habitual behaviour the agent doesn’t know “what counts” as going on in the same way. Since Brett (1981) effectively argues against both writers in rejecting the caricatures of habit with which both writers are working, I do not rehearse those arguments here.

direct control of the agent, and as we are often reminded, control and responsibility come and go together.

The sense of control which distinguishes the two sorts of behaviour is the agent's ability to directly intervene on their performance. Normally one cannot intervene on reflexes or bodily processes. There are exceptions that need qualifying: if one anticipates a reflex, such as a knee jerk, one can attempt to stop it. But that is not achieved by a direct act of will, but by doing something else to inhibit the reflex, such as tensing muscles. And whilst one might be said to have the capacity to intervene on some bodily processes such as breathing, the possibilities for intervention are limited: one can affect the way one breathes within certain limits, but not whether one does.

But intervention in habitual behaviour can always be direct and comprehensive. By an act of will, one can stop oneself from exercising a given habit, or modify it to a greater or lesser extent, on any particular occasion, merely by paying attention to what one is doing. It is because of this that we can say that habitual behaviours are within the scope of agency, in a way in which reflexes and bodily processes are not.

There are other classes of automatic behaviours with which habits might be confused to negative effect. Indeed there may be many such classes, but let us focus on three that are commonly mentioned, in the hope that what can be said about these classes will be applicable quite generally. These are addictions, compulsions and phobias. Examples of each are alcoholism, compulsive hand-washing, and arachnophobia respectively. Confusion with habits might arise because ordinary language allows that all of these conditions can be spoken of, quite correctly, as "habits": we talk about alcoholism as a "bad habit", and there seems no error in speaking of compulsive hand-washing as a "habit" that is taking over somebody's life, or somebody's propensity to run from spiders as a "habit" that goes back to a bad childhood experience. Speaking of these behaviours in these terms doubtless draws attention to what they share with habits, in particular, their being repeated and automatic.

The worry, however, is that when we use any of these three conditions to explain an agent's behaviour, we do not imply that the agent is fully responsible for them, or at least not in a direct way. On the contrary. Addictions, compulsions and phobias are all "conditions" that the agent "suffers", and their manifestations are accordingly not actions in any full sense.

For instance, when we explain somebody's having a drink by referring to their addiction to alcohol, we imply that there is some kind of physiological dependency which is driving the behaviour, which is external to her agency. In an important sense, we are saying that she can't help herself. That, after all, is one of the worrying things about addiction (along with the damage that repeated consumption causes). Precisely how much control the addict lacks over her addictive behaviour will vary from addict to addict, some being able to avoid opportunities to buy drink, others having to attend AA meetings to do even this. But it is doubtful that we should ever say that addiction is as external to agency as a reflex. After all, it takes some complex intentional moves to buy a drink.

Of course an addict might be said to be indirectly responsible for becoming addicted, since she freely chose to have the drinks at the outset, she knew the risks, and to that extent she is responsible for having this drink. But when we explain a piece of behaviour in terms of an addiction, we are not making any claim about how she became addicted. Now she is addicted, her addictive behaviour falls short of action, and that is what we imply when we explain behaviour with addictions.

Analogous things can be said about compulsions and phobias. When we explain someone's behaviour in terms of a compulsion, we attribute to them a psychological condition whose manifestations get in the way of normal life, and are contrary to what the agent herself judges it best to do. And when we explain somebody's behaviour by means of a phobia, we attribute a condition which is triggered by the presence of certain sorts of thing (in the example, spiders), and towards which the phobic reaction is inappropriate, often by the agent's own lights. As with addictions, behaviours explained in terms of compulsions and phobias fall short of full agency.

Contrast these cases with a piece of behaviour explained with a habit. When we do this, the suggestion is, we do not imply that responsibility is in any way reduced. Consider, for instance, Alasdair's explaining his taking a particular route between the station and his office this morning by saying that it his habit to take that route. He does not thereby try to reduce his responsibility for doing it. And this would be the case no matter how he acquired the habit. If he wanted to escape responsibility for his action, he would not use the term "habit" in his explanation. That at least is the suggestion. In the next section we shall see some positive reasons for relating habits to agency in this way.

## 2. HABITS AND AGENCY

It will be suggested in this section that there are two main factors that support the common-sense intuition that exercises of habits are manifestations of agency. What is striking is that neither of these factors rely on the agent being implicated in some prior psychological state or event. And that has implications for how we ultimately understand the nature of habit explanation.

First, if somebody does something out of habit, it is perfectly within their power to have done otherwise were they to choose to do so. They have what has been called *intervention control* over their habitual behaviours.<sup>3</sup> This is to say that the agent could intervene at any point before or during the behaviour in question, and do something else, or nothing at all. The suggestion is that since it is the agent herself who has this sort of control over the exercises of her habits, it is she who is responsible for them. It would therefore be rational to hold various attitudes towards her, say of blame or praise, since she could have done otherwise.

Intervention control is notably lacking in manifestations of the three sorts of condition mentioned above. The alcoholic, for instance, cannot intervene in her drink buying, or at least not with any reliability, and that lack of intervention control is what compromises her responsibility for doing so.

Of course, as Harry Frankfurt (1969) has famously argued, the possibility of doing otherwise may not be necessary for somebody to be responsible for their behaviour, but it seems to be close to being sufficient. It is not quite sufficient because it may be possible for somebody to intervene in an activity which does not appear to be an action at all, say the behaviour of some inanimate object. Consider for example Frankfurt's (1978) coasting car, with the driver not touching any controls. It might also be possible to intervene in somebody else's action, as for example when a driving instructor has dual controls. But in neither case does intervention control make the intervener properly "the author" of those activities, as we might put it. She may bear *some* responsibility for these things going right, but that is not enough for her to be fully responsible for them. It is the second consideration concerning habits which does seem to secure authorship, by bringing responsibility up to this level.

---

<sup>3</sup> See also Pollard (2003, pp. 415-6).

The second consideration which supports the intuition that habit explanations convey full responsibility is the intimate connection habits have to agency. The thought here is that having a certain habit is part of what it is to be a particular agent. This can be most easily seen when we consider the process of habit acquisition, which is analogous to the movement from novice to competence in a given technique.

Before an agent acquires a habit of  $\Phi$ -ing, she is able to  $\Phi$ , though cannot yet do so automatically.  $\Phi$ -ing requires thought or concentration, though these efforts need not be explicitly directed towards acquiring the habit, since one can acquire a habit inadvertently. At this stage, the agent as it were goes through the motions of somebody who has the habit, since she does not yet do it automatically; it is not yet “second nature” to her. She has not made  $\Phi$ -ing her own.

Given a number of repetitions, the test of whether  $\Phi$ -ing has become a habit is not only that  $\Phi$ -ing has become part of her history, but also for  $\Phi$ -ing (in these circumstances) to be automatic. She no longer needs to think about doing it; indeed it is easier for her to do it than not to do it. It is at this point that it makes sense to say that the agent has made the habit of  $\Phi$ -ing her own.  $\Phi$ -ing (in characteristic circumstances) is something that *she* now characteristically does. The suggestion is that we would now be leaving something out of a description of what it was to be that agent were we to omit her habit of  $\Phi$ -ing.  $\Phi$ -ing is now not only second nature, but it is part of *her* nature. Exercising the habit is part of who she is. The same thing could not have been said before the habit was properly acquired.

William James once suggested that “the young” will become “mere walking bundles of habits” (James 1891, p. 127). But it would be hasty to say this of the young or of anybody else for that matter. For there will be things besides habits that will combine with them to make up agency. Presumably agency will also include many psychological and physiological features, all of which will be part of a complete account. But, the present suggestion is that an account of agency which focuses only on these latter things will have missed something important out. For habits are an essential part of the picture. Many habits will be essential to agency as such (such as those to do with language and reasoning), others will be, as it were, optional, and could vary from individual to individual (such as nail biting and mannerisms). But all of an agent’s habits contribute to making that individual who she is.

If this is right, explaining a behaviour by making reference to one of the agent’s habits brings authorship with it for free. If a given piece of behaviour is explained by the

agent's habit, and her habit is part of who she, the agent, is, then it follows that the behaviour must be hers. She is its author.

Thus it seems that the intervention control we have over exercises of habits, plus the constitutive relation habits bear to agency, are together sufficient for us to say that the exercise of a habit is an action. What is more, it has been possible to do this without implicating psychology in the actual (i.e. non-counterfactual) performance of habits. That is not only a considerable explanatory advance, leaving no need to give an account of the relation between the mental and the bodily, but it also suggests the insistence on psychology in action explanation is an intellectualist prejudice.

Finally, we can return to the confusion of §1, namely, the observation that ordinary language permits us to speak of addictions, compulsions and phobias as habits. We can now say that insofar as we are prepared to accept an explanation in terms of addiction, compulsion or phobia, we are acknowledging a diminution of responsibility on the part of the agent. But insofar as we are prepared to accept an explanation in terms of habit, we are not doing this, but rather drawing attention to an aspect of that agent, as she conducts herself in a fully responsible way. We would do well - both as philosophers and as folk-explainers - to remind ourselves of these implications when we are choosing our explanatory terms.

### 3. INTENTIONALITY

There is another possible objection to the claim that habits explain actions proper which, given the orthodoxy in the philosophy of action, must be addressed. The problem concerns whether (so called) habitual actions are intentional under any description. For if they are not, and Davidson (1980) is right to say that every action is intentional under *some* description, it follows that habitual behaviours (as we should have to call them) may not be actions at all. How are we to respond to this objection?

First, there does not seem to be too serious a difficulty with intentionality when habitual behaviours are performed as part of some larger intentional activity, whatever the grounds for attributing intentionality to that activity may be. An example is when one habitually puts on one's socks as part of getting dressed. As long as we have grounds for saying that getting dressed is done intentionally, we can say with Brand (1984, pp. 28-30), that any habitual actions which are involved in that activity "inherit" the larger activity's intentionality.

But this cannot be part of a general answer since we can exercise habits when they are not part of a larger activity, as well as when they are part of some larger activity that is



unintentional. As an example of the first sort, consider habitually biting one's nails. Now this is not an unintentional action in the classic sense that knocking over the glass whilst intentionally reaching for the salt is unintentional. For there is nothing else one is doing intentionally in the course of which biting one's nails is merely an accidental upshot. And there does not seem to be an ascribable intention (whether or not that is understood as a belief and a desire) which might be said to make one's nail biting intentional. One typically will not have intended to do it, and will have no positive attitude towards it. So the actional status of the behaviour is in doubt.

Matters are no better with examples of the second sort. Consider the example (from the introduction) of somebody missing a turning because she habitually goes straight on. It may look as though missing the turning is an unintentional action. But again, using standard resources, this would be wrong. For she did not go straight on intentionally. And this is borne out if we acknowledge that there is a minimal consistency requirement between what one intends and what one does intentionally.<sup>4</sup> For it seems that what makes sense of her "missing" the turning is that she must have had an intention, presumably formed at an earlier time, to take it, or at least to do something that required taking it. And consistency requires that she could not also intentionally miss it. So in this second sort of case we have stronger reason than before to deny that there is a description under which this putative action is intentional.

It nevertheless seems wrong to say that the sorts of behaviour described above are not actions. But what grounds can be given for this intuition? It would be no answer to simply relax the Davidsonian requirement that actions be intentional under some description. Not only would that be to abandon a well established truth in philosophy of action, but more to the point, we would still lack a positive reason for considering the actions in question to be intentional.

One possibility is to suppose that the intentionality of habitual actions is conferred by what Michael Bratman (1984) calls "present-directed intentions". These states are supposed to be contemporaneous with the action and, given a suitable content, are supposed to account for the intentionality of actions quite generally. However, they seem ill-suited to the habitual case. The fundamental problem is that in accordance with what Bratman calls the

---

<sup>4</sup> Such a requirement is entailed by all of the views considered by Bratman (1984).

*“methodological priority of future-directed intention”* (1984, p. 379), present-directed intentions are modelled on future-directed intentions. This might enjoy some plausibility as well as some explanatory benefits.<sup>5</sup> After all, we can make sense of the idea of more or less time between forming an intention and acting upon it, so why not contemporaneous intentions? And this move allows us to model the normative relations that must stand between intentions and intentional actions of the sort referred to above, on those that obtain between intentions and other states. But if our focus is on habitual behaviours, the idea of temporal leeway makes no sense. For what is striking about all habitual behaviours is that they do not require prior thought or planning, and nor does thought or planning ensure one performs them. It is repetition – habituation – that is required. And that is not a cerebral matter, but a practical one. So any account of intentionality modelled on deliberative processes merely introduces an intellectualist distortion.

There is a more promising possibility that does not presuppose that intentionality must be understood in an intellectualist way. The idea is to say that the exercises of habits have some sort of “intrinsic” intentionality. Specifically, it is to say that such behaviours have a certain teleological structure – they have a point, purpose or goal. In the above examples, these would be that one’s nail are bitten, and that one gets home in the normal way respectively. These goals could be made explicit, but should not to be understood in terms of some prior item which somehow bestows intentionality on the otherwise non-intentional behaviour. Nor indeed need such goals be understood as being the same as the agent’s reasons for acting – for the agent need have no conception of them at the time of acting – though she might be brought to accept them in due course.<sup>6</sup>

This idea can still accommodate the normative constraints stressed by Bratman, such as requiring the intrinsic goal not be inconsistent with other goals that the agent may have (perhaps in the form of intentions). This would allow us to see what is wrong with the missing the turning case. But there need be no suggestion that any such goals are within the agent’s intellectual grasp at the time of acting.

---

<sup>5</sup> Doubts are raised by Velleman (1991, pp. 278).

<sup>6</sup> See Pollard (2003) for a suggestion of this sort.

There is of course much more to be said about this intrinsic notion of intentionality.<sup>7</sup> But the above should suffice to show that there is conceptual space for a notion of intentionality which could apply to habitual actions, which would allow us to agree with Davidson that all actions are intentional under some description, without being forced to intellectualise habitual actions by positing unwanted psychological apparatus.

#### 4. THREE ACCOUNTS OF THE EXPLANATORY RELATION

We now turn to the question: What sort of explanation is a habit explanation? To answer it let us consider three sorts of explanation which can be discerned in the literature on action explanation, though they may not be mutually exclusive. It will be argued that none properly capture the idea of habit explanation. This motivates the introduction, in the next section, of a fourth kind of explanation.

First, a habit explanation is not a psychological explanation. Such explanations explain by mentioning some item or items in the agent's psychology, that is, items which exhibit the characteristic first-third person asymmetry. This is a very large class of explanations whose *explanantia* range from tickles and pains on the one hand, to thoughts and attitudes with specific propositional contents on the other. In giving psychological explanations we might also be giving some sort of justification of the action from the agent's point of view: what Davidson (1980) calls "rationalizing". But equally, justification might not be at issue, as when one explains one's jumping by referring to the pain one felt.

Habits are not psychological items in this sense. There is no distinctive feel to having a habit, as there is for pains and tickles. And the agent is not even authoritative about whether they have a particular habit. After all, many habits are acquired quite inadvertently, and we often don't notice whether we have them or not, even if we exercise them on a daily basis. One may for instance have a habit of putting one's left sock on before one's right sock,

---

<sup>7</sup> For instance, the idea of intrinsic intentionality bears comparison with John Searle's notion of "intention in action" Searle (1983, Ch. 3). Searle's notion is motivated by the failure to locate a prior intention for "subordinate" or "sudden" actions. But habitual actions need not be of either kind. As we saw above, there may be no larger activity of which they are part, and habitual actions can be performed slowly and carefully, just as taking the same route to work might be. Indeed another advantage of conceiving of intentionality in this way is that it allows us to make sense of the many intelligent adjustments that might be made in pursuing the goal in question.

but have never noticed it. Indeed, unless habits are brought to our attention, we typically do not know we have them at all. This contrasts with the perspective not merely of any other person who might be a stranger to us, but of those with whom we share our lives, our spouses, parents, colleagues or friends. It will typically be these people who notice our habits, and they will be the ones who bring them to our attention if anybody does, especially the annoying ones. Epistemologically, then, habits enjoy what has been called *second-person authority*.<sup>8</sup>

Second, habit explanations are not teleological.<sup>9</sup> They do not specify the point, purpose or goal of the action being explained. That is not to say that the action being explained does not also have a teleological explanation (as we saw in §3), nor indeed that the habit itself might have such a teleology. An example of the first would be somebody's going to the bank to get some money out of habit. Here, each exercise of the habit has a specific purpose. But we are in possession of a hybrid explanation that specifies both teleology and habituality, and we could have one without the other. An example of the second would be somebody's habit of running every day in order to keep fit. An exercise of that habit might be said to be teleologically explained by referring to the aim of fitness. But again, this would be a hybrid explanation. There is reason to think that we would still be giving something of an explanation of today's running by referring to the habit alone, and not to the intended goal. Though that of course would be leaving much unsaid.

The third possibility is that habit explanations are a variety of causal explanations.<sup>10</sup> The fact that we sometimes use the expression "force of habit" in habit explanations certainly seems to lend support to this view. What is more, causal interpretations enjoy currency in philosophy not least because they are well understood by science. If habits were causes, perhaps in the same way as the physical dispositions standardly discussed by philosophers, such as brittleness and elasticity are causes, a naturalistic picture of action could be significantly advanced. But there are reasons to doubt that habits can be causes.

Part of the problem is that habits are not like the standard sorts of dispositions mentioned above. As philosophers writing on such matters are fond of pointing out,

---

<sup>8</sup> Pollard (2005, p. 80).

<sup>9</sup> Examples of teleological explanatory accounts include Collins (1987) and Stout (1996).

<sup>10</sup> Davidson (1963) is the *locus classicus* of this view. See also Goldman (1970).

standard dispositions need not have been manifested by the sample in question, for those properties to be correctly attributed to that sample. This fact supposedly generates a certain problem about how to distinguish between two “untested particulars”, one of which has the disposition, whilst the other lacks it. And this problem is standardly given a metaphysical solution, according to which the disposition is ultimately constituted by the sample’s microstructure, or some metaphysical analogue.<sup>11</sup>

However, when we look at habits, we see that a person with a particular habit is logically required to have a certain sort of history. If somebody has never  $\Phi$ -ed before, it is simply false to say that she has the habit of  $\Phi$ -ing. Thus in the case of habits, there is no such thing as an “untested particular”, and hence no problem of how to distinguish between two of them. An agent’s history of doing something, plus her present proficiency at it, are together sufficient for us to decide whether she has the habit of doing it or not. Whether or not metaphysics is required to analyse the standard dispositions, it is not required here.

But even if we accept that habits are not like standard dispositions on the grounds that they must have manifestations, there is still a problem for the causal interpretation. For those manifestations upon which the very existence of the habit depends cannot themselves be caused by the habit, because they are conditions for the existence of that habit. Causal explanations are supposed to be asymmetrical in the sense that they support counterfactuals of the sort “had the cause not existed, the effect would not have existed”, and not the other way round. But in the present case we do have a counterfactual running in the opposite direction: had the supposed effect, the action, not existed, that particular habit would not have existed.

Of course, if one was determined to force habit explanation into the mould of causal explanation, one might posit some common cause responsible for all of the manifestations of a particular habit, and suppose that this cause is identical with the “habit” referred to in a habit explanation. But this would require such a serious revision of our ordinary concept of habit that it could not possibly provide a satisfactory analysis of the habit explanations we are familiar with.

---

<sup>11</sup> Mumford (1998, p. 29) for instance, proposes a “realist” analysis of dispositions, which posits a “categorical base” which one sample has, and the other lacks.

This is not to deny that habitual actions and habits themselves are dependent upon causal processes. On the contrary, they are totally dependent on many such processes working correctly, be they processes in the perceptual system, the brain or the muscles. But what it is to deny is that when we explain an action in terms of a habit, we are referring to any such processes.

An important motivation for persisting with any causal interpretation is the distinct lack of alternatives. And the above considerations do suggest a different understanding of the explanatory relation, which is seldom recognised, but is intuitive and appealing in the present case. This is *constitutive* explanation.

## 5. CONSTITUTIVE EXPLANATION

Constitutive explanation is the sort of explanation we give when we explain something by saying that it is a part of – constitutes – something else. Examples would include explaining what a particular cell is by saying that it is part of a larger organism; explaining what a broom-handle is by saying it is part of a broom; or explaining Bob's attendance at a lecture by saying it is part of his degree course. The proposal is that when we give a habit explanation we are doing something of the same sort. We are explaining what an action is by saying it is part of somebody's habit.

First some remarks on the metaphysics of constitution. Constitution is here to be distinguished from identity. Some writers defend a view according to which if a plant is constituted by the cells that make it up, then the plant is identical to those cells.<sup>12</sup> If this were right we would be committed to saying that a habit is identical to the totality of its exercises, and this would be problematic not least because it makes it hard to see how a habit could be explanatory. If a habit were really nothing but the totality of its exercises, the similarity of the exercises might be purely accidental, and hence could not really explain anything. But the relevant relation of constitution allows that an entity can be said to be composed of other entities without remainder, but that entity is still distinct from the totality of those entities.

There are two main considerations. First, a constituted entity is not constituted by any particular set of constituting entities. So a plant is not constituted by any particular collection of cells – each of them might have been another similar cell. And a habit is not

---

<sup>12</sup> For example Burke (1992) and Olson (2001).

constituted by any particular set of actions – each action might have been another similar action. Second, a constituted entity can have properties and stand in relations that the totality cannot. A plant can be healthy, or belong to me, and a habit can be hard to break, or irritating. Plants and habits, then, are genuine entities in their own rights.

An objector might point out that whilst this idea of constitution might apply to plants and brooms, it cannot work for habits as stated, since it cannot capture the persistence of habits, that is their continued existence whilst not being exercised. The problem is that the constituting entities, the actions, are only present at particular times, and are absent in between. So if a habit is constituted by its exercises, it seems that we are forced to say that the habit only exists when its exercises exist. And we should resist any account which has this consequence.

But the problem disappears providing we distinguish between constitution *at a time* from constitution *across time*. The former applies to plants, brooms, and other composite objects. The latter applies to entities whose constituents are spread across time, and may not be contiguous, like somebody's doing a philosophy degree, building a house, or (according to the present proposal) having a habit. We could thus correctly say that Bob's doing a philosophy degree consists in activities such as attendance at lectures, writing of essays, revision, and exams. Doing the degree is made up of these distinct activities, and if he didn't do these things, he would not be doing a philosophy degree at all. What is more, we can time the beginning of his doing the degree as simultaneous with Bob's involvement with the first constituting activity, and its ending as simultaneous with the last. What we would not be forced to do therefore, is deny that when Bob is having a drink, visiting home, or sleeping, that he is still doing his degree. For whilst these are not constituents of the larger activity, they take place within the timeframe of that larger activity. So constitution across time works differently from constitution at a time. If the constitution of habits is conceived in this way there is no difficulty in saying both that a habit is constituted by its exercises, and that it persists even when it is not being exercised.

What then of counterfactual dependencies? First of all it should be clear that interpreting habit explanations as constitutive does not undermine the obvious fact that they support counterfactuals that run in the usual direction, that is, from *explanandum* (the action) to *explanans* (the habit). Had the agent not had the habit, she would not have done such and such action is still made true by saying that the action partly constitutes the habit. For in

saying this we are naming one of the formal properties of the action, namely its constitutive relation to the habit.

What of counterfactuals in the opposite direction? In §4 we saw that an objection to the causal interpretation was that habit explanations support counterfactuals that run from *explanans* (the habit) to *explanandum* (the action), and that is the wrong direction for a causal explanation. But if we interpret the explanatory relation as constitutive, that such counterfactuals are supported is no objection. For since a constitutive explanation articulates only formal properties of the *explanandum*, there need not be any prohibition on such properties being equally illuminating in the other direction (though since the properties are complementary, one would not normally require both). Indeed it would seem to be a strength of such explanations that they tell us something not only about the *explanandum*, but also about the *explanans*.

For example, when one is told in a constitutive explanation, that cell C partly constitutes plant P, one learns not only that cell C is part of plant P, but also that plant P has cell C as a constituent. Two counterfactuals are thus supported: had plant P not existed, cell C would not have existed, or not at least in the same way; and had cell C not existed, plant P would not have existed, or not at least in the same way.<sup>13</sup>

Transposing this for habit explanation, from the explanation of action  $\Phi$  by habit H, we get: (i) had habit H not existed, action  $\Phi$  would not have existed, or not at least in the same way; and (ii) had action  $\Phi$  not existed, habit H would not have existed, or not at least in the same way. Both counterfactuals are supported by a habit explanation as long as we understand it constitutively.

What is more, both (i) and (ii) seem to be true on independent grounds. What makes (i) true is that had the agent not had a certain habit she would not have done quite what she did – she might have done something similar, say, deliberated and then  $\Phi$ -ed, but that would not have been quite the same action, for it would have lacked the automaticity of habitual action. What makes (ii) true is that had an agent not  $\Phi$ -ed, the habit H would not be quite the habit it is; in particular it would not have been further entrenched (perhaps “reinforced”)

---

<sup>13</sup> The plant-cell example is merely an illustration; the reader who has problems with the example should substitute a less problematic composite object.



by the agent's present  $\Phi$ -ing. These familiar truths about habits further support to the view that habit explanation is constitutive.

The proposal, then, is that when we give a habit explanation, we are not explaining the action in terms of psychology, teleology, or causes, but in terms of one of its formal properties, namely, that portion of the agent's career of which that action forms a part. That may not be all we convey when we give a habit explanation: we may for instance be saying something essentially negative, to the effect that deliberation is not involved. But conceiving the core of such explanations as constitutive allows us to see how they convey something positive about their relation to the agent.

If this is right, the availability of constitutive explanation also tells us something philosophically important about the structure of action as such. If we are interested in giving anything that aspires to be a complete account, this will have to be part of it.

## 6. HABITS IN THE PHILOSOPHY OF ACTION

The proposal is that not only can habits explain actions, but also that they explain them in a distinctive, that is constitutive, way. This has a number of implications as well as benefits for the philosophy of action.

Amongst the implications are that we now have an expanded conception of what it could mean to say that we have "explained" an action. Hitherto we could say that we have explained an action if we knew what caused it, what it aimed at, what the agent was thinking or feeling, or perhaps a combination of these. Now we can also say that we have explained an action if we have said how it fits in with her having done this sort of thing before. In other words, we have located the action in one pattern in the agent's career. This is the respectable face of behaviourism, according to which behaviour is an important, but not the exclusive, form of explanation.

This expansion of explanatory possibilities points to a second implication. For if the above argument is right, we have also expanded our conception of what it means to be an agent. Namely, not as an entity narrowly conceived as the locus of psychological states and events which alone give rise to actions, but rather as an entity whose identity is inseparable from its familiar modes of behaviour, in its familiar environment, which stretches back in time as well as forward. Agency, then, to adopt some fashionable jargon, is conceived as both "embodied" and "embedded".

When we consider just how much of our lives we spend exercising habits, rather than subjecting our actions to deliberation, what is suggested is a research programme for the philosophy of action which no longer places psychology (beliefs, desires, intentions, reasons) at its centre, but rather at its margins. For whilst thought is very helpful when we are in novel or important circumstances, the rest of the time it rather gets in the way. In a slogan we might say: we only think when our habits give out.

*University of Edinburgh*

## REFERENCES

- Brand, Myles. 1984. *Intending and Acting*. Cambridge, Massachusetts: MIT Press.
- Bratman, Michael. 1984. "Two Faces of Intention," *Philosophical Review*, 93, pp. 375-405.
- Brett, Nathan 1981. "Human Habits," *Canadian Journal of Philosophy*, 11, pp. 357-76.
- Burke, Michael B. 1992. "Copper Statues and Pieces of Copper," *Analysis*, 52, pp. 12-17.
- Collins, Arthur C. 1987. *The Nature of Mental Things*. Notre Dame, Indiana: University of Notre Dame Press.
- Davidson, Donald. 1963:. "Actions, Reasons, and Cause," in Davidson (1980), pp. 3-21.
- 1980. *Essays on Actions and Events*. Oxford: Clarendon Press.
- Frankfurt, Harry. 1969. "Alternate Possibilities and Moral Responsibility," in Frankfurt (1988), pp. 1-10.
- 1978. "The Problem of Action," in Frankfurt (1988), pp. 69-79.
- 1988. *The Importance of What We Care About*. Cambridge: Cambridge University Press.
- Goldman, Alvin. 1970. *A Theory of Human Action*. Prentice Hall.
- James, William. 1891. *Principles of Psychology*. Macmillan.
- Mumford, Stephen. 1998. *Dispositions*. Oxford: Oxford University Press.
- Olson, Eric T. 2001. "Material Coincidence and the Indiscernability Problem," *Philosophical Quarterly*, 51, 337-55.
- Pollard, Bill. 2003. "Can Virtuous Actions be Both Habitual and Rational?" *Ethical Theory and Moral Practice*, 6, pp. 411-25.
- 2005. "Naturalizing the Space of Reasons," *International Journal of Philosophical Studies*, 13, pp. 69-82.
- Ryle, Gilbert. 1949. *The Concept of Mind*. Penguin Books.
- Searle, John R. 1983. *Intentionality*. Cambridge: Cambridge University Press.
- Stout, Rowland. 1996. *Things That Happen Because They Should: A Teleological Approach to Action*. Oxford: Oxford University Press.
- Velleman, J. David. 1991. "Book Review: Intention, Plans, and Practical Reason," *Philosophical Review*, 50, pp. 277-84.

Winch, Peter. 1958. *The Idea of a Social Science and its Relation to Philosophy*. London and Henley: Routledge & Kegan Paul.