



## BIROn - Birkbeck Institutional Research Online

Hornsby, Jennifer (2008) Agency and alienation. In: de Caro, M. and MacArthur, D. (eds.) *Naturalism In Question*. Cambridge, U.S.: Harvard University Press, pp. 173-187. ISBN 9780674030411.

Downloaded from: <https://eprints.bbk.ac.uk/id/eprint/96/>

*Usage Guidelines:*

Please refer to usage guidelines at <https://eprints.bbk.ac.uk/policies.html>  
contact [lib-eprints@bbk.ac.uk](mailto:lib-eprints@bbk.ac.uk).

or alternatively

## Agency and Alienation

**1.** David Velleman's 'problem of agency', in his paper 'What happens when someone acts?', is the problem 'of finding an agent at work amid the workings of the mind' (1992, p.131). This problem arises when 'a naturalistic conception of explanation', implicit in the 'standard causal story' of action as belief-and-desire-caused-behaviour, is adopted.<sup>1</sup> In my opinion, the standard story, as it is standardly naturalistically understood, should be rejected. Rather than seeking an agent amidst the workings of the mind, we need to recognize an agent's place in the world she inhabits. In order to do so we have to resist the naturalistic assumptions of the standard causal story.

I use 'naturalism' and cognates here in Velleman's sense, which is the sense of my opponents. (Thus 'naturalism' stands for a doctrine rejected by many authors in the present anthology, albeit that we may subscribe to a different doctrine at least as deserving of the name naturalism.) I single out Velleman's 1992 paper for criticism because Velleman is very explicit there about his naturalistic prejudices, so that it provides a clear example of a certain style of thinking.<sup>2</sup>

**2.** There are various phenomena of human agency which we are apt to describe using the language of alienation, or of estrangement, or of non-participation. In Velleman's account, the agents who are alienated are those who lack self-control, or self-understanding, or who undertake projects half-heartedly. There is the addict who injects heroin in spite of knowing that it would be better if she could resist doing so. There is the person who lacks motivation, because of depression or fatigue. And there is the agent who finds herself accounting for her behaviour by saying 'It was my resentment speaking not I'. In all of these cases, Velleman thinks that we have belief-desire caused behaviour, and that the standard causal story applies. But the standard story is adequate to these cases, Velleman thinks, only because in these cases the relation between a human being and her action falls short of what is needed for a case of genuine agency. That is why Velleman tells us that we have to embellish the standard story to characterize what he calls agency *par excellence*: that we can only characterise the relation that obtains between a human being and her action in cases of *non-alienated* agency if we add something extra to the standard story's states and events. But my idea is that Velleman's problem of agency vanishes when the standard story is discarded.

I want to discuss different kinds of alienated agency in what follows, in order to try to corroborate my opinion that the standard story should be rejected, not embellished. Besides the real phenomena that are describable using the language of alienation, there is alienation of a kind that I shall call 'unthinkable'. We know

-----  
1 The 'standard causal story' of action is widely credited to Davidson. Certainly Davidson's work has done a great deal to ensure that the thesis that explanation of what people do which proceeds by giving their reasons is *causal* explanation; and a certain understanding of this thesis gives rise to the standard story (see §6). But I think that Davidson's claims of the mental's irreducibility ought to discourage the picture of action that I criticise here (and that it would discourage it if a particular view of causality, also Davidsonian, weren't in play). There is further elaboration of the standard story in §§5 and 6.

<sup>2</sup> Velleman 2000 may not be such an obvious target for my remarks, as I acknowledge in an Afterword. Velleman is not alone in allowing his naturalistic thinking to lead to very implausible accounts of ourselves as agents. In my forthcoming, I pick on Michael Bratman.

that we are not alienated agents of the unthinkable kind, so that we can also know that we must resist whatever assumptions lead to our feeling that we might be so alienated, and that we must not assimilate the real phenomena of alienation to the unthinkable kind. I shall use a discussion of unthinkable alienation (§§3, 4) to elicit the errors of the standard causal story of action (§6) and to show that Velleman's problem is misconceived (§7). That will put me in a position to say how we might think about phenomena which interest Velleman and which we describe using the language of alienation (§8).

**3.** When he endorsed the conception of explanation that gives rise to his problem of agency, Velleman cited Thomas Nagel. Nagel invited us to adopt a picture of the world in which all events and states of affairs are seen as caused either by other events and states or by nothing at all. If you try to imagine your actions as part of the flux of events in this picture, then you will find yourself alienated from them. As Nagel put it:

Everything I do or that anyone else does is part of a larger course of events that no one "does" but that happens.' (1986, p.113).

It seems that in order to adopt the picture which Nagel invites us to, we have to view actions as set apart from the agents whose actions they are. But if actions are events—as Nagel assumed in presenting the picture, and as my naturalist opponents here all assume—then they are surely not events from which agents are set apart. The phenomenon of human agency can be caught in the first instance with the idea of *someone's doing something intentionally*. When that idea is put together with acceptance of an event ontology, there is a way to define 'an action'. Thus, one might say: there are human beings who do things; when someone's doing something is her doing something or other *intentionally*, human agency is exemplified, and an event that is *her doing the thing* is an action. It is no wonder then that we should feel alienated if we are meant to think of our actions among the course of events and proceed to speculate about how we might fit in. For where an action has been picked out, an agent has been: the action is *her* doing something. (It isn't true quite in general that an agent has been picked out whenever an action has been. One might for instance know that some human being had caused something without knowing who had. Such examples, however, do nothing to suggest that one ought to look for a human being within the flux of events present in the picture from Nagel's external perspective.)

Nagel encourages a sense of alienation by speaking as if you stood to an event that is your action in a relation expressible using the word 'do'. This makes it seem as if you could participate as agent only by being related to something that might be present in a scene in which you yourself were not involved. It can then be tempting to think that in order to make a difference, you would have to butt in as a cause at the point at which your action is found. Hence, perhaps, some of the attractions of the claim that agents cause actions. But even if the idea of agents' butting in might somehow help to give sense to the thought that agents contribute to what happens, there would still be a difficulty about supposing that the world to which agents make a difference is occupied only by events and states that are part of a flux in which agents themselves might never have been involved. For we take ourselves be influenced by, not only to act upon, the world to which our actions make a difference. In order to escape from a general threat that we are alienated from the world we inhabit as agents, we have to avoid thinking of ourselves as standing in a relation either of *doing* or of *causing* to the events that are our actions.

(It should be acknowledged that it is only in a semi-technical, philosophical usage that ‘action’ stands for events, and that it is not ordinarily so used. But it is only in this sense that actions are particulars and thus candidates to be ‘part of a larger course of events’. It adds to the general confusion in this area that ‘event’, as well as ‘action’, has uses in which it doesn’t stand for particulars. My policy here is to use both ‘action’ and ‘event’ as I think naturalists mean to—only for things in an ontology of particulars.)

When the relation between agents and the events that are their actions is understood, it will not seem possible to locate actions among a causal flux in which agents might play no role. Human agents are not merely things within which things happen, and they clearly do play a role in the arena within which their actions are found. For an event of someone’s doing something is typically an event of her bringing something about; and the event that is her action (her doing the thing) brings about that which she brings about. A driver slams on her brakes and brings it about that the car comes to a sudden stop; the event that is her slamming on the brakes brings it about that the car come to a sudden stop.

Nagel was surely right, then, to say that the very idea of agency is threatened when we try to accommodate actions in an ‘external perspective’. The role of agents in a world of events is evident only when it is appreciated that agents cause things—things that ensue from their actions. It seems unthinkable that agency should be manifest from any point of view from which it is impossible to locate agents.

**4.** Nagel asked us to imagine looking at things from far away in order to pose his question about actions’ place in the natural world. One can produce the mystery also by looking at human beings close up and looking inwards—as Hume did.

Hume expressed great bafflement about our role in the explanatory order when he looked at agents’ insides:

We learn from anatomy that the immediate object of power in voluntary motion is not the member itself which is moved, but certain muscles and nerves and animal spirits, and, perhaps, something still more minute and more unknown, through which the motion is successively propagated ere it reach the member itself whose motion is the immediate object of volition. ... [T]he power by which this whole operation is performed, so far from being directly and fully known to an inward sentiment or consciousness, is to the last degree mysterious and unintelligible. How indeed can we be conscious of a power to move our limbs when *we have no such power*, but *only* that to move certain animal spirits which, though they produce at last the motion of our limbs, yet operate in such a manner as is wholly beyond our comprehension? (1748, §7, Pt I.)

There are two things to notice about this passage. (Hume’s own agenda here are not to the present point.<sup>3</sup>) First, Hume’s proffered reason for denying that we are conscious of a power to move our limbs is that we do not have such a power.

-----

<sup>3</sup> At this stage in the *Enquiry*, Hume is in the process of arguing that we have no impression of causal power. He aims to refute someone who says that such an impression can be got from reflection on the influence of our volitions on bits of our bodies that we can move.

Secondly, Hume's denial that we have a power to move our limbs does not (at least at this juncture) spring from any general mistrust of the idea of power; for he tells us straightaway that there *is* a power we have—a power to move certain animal spirits, whose excitations eventually produce motions of limbs. Hume's is then a view of ourselves as agents from which we are bound to feel estranged. The only power we have is a power to produce effects, which are, as Hume says, 'totally different from' the ones that we intend. Such effects—events in brains—not only fail to be within the scope of our intentions, but seem not to belong to the world that we inhabit as agents. Agents, in a word, are alienated.

One might want to blame Hume's strange view on his ignorance of science, and one might want to blame it on his dualistic thinking. But I suggest that it has another source. If one looks inside someone's skull, expecting to see the makings of intentional bodily movements there, one is sure to encounter a mystery about them.

To see this, imagine someone called Jane who is given a very detailed account of electrochemical impulses, neural transmission, and so on, but an account that doesn't mention any organism inside which this all takes place. Jane is told that, by knowing the scientific story, she knows *everything* that happens when someone moves his arm at will. She might reasonably be puzzled. Certainly knowledge of relevant portions of electrochemistry and understanding of the operation of neural transmission make her better placed than Hume was to know what goes on when someone moves: she can understand how a limb comes to move. But science doesn't make her better placed than Hume was to say what someone's intentionally moving his arm *consists* in. Hume's mystery does not go away by providing a neuroscientific account of what goes on inside.

Philosophers who are used to thinking that accounts at the sub-personal level record all the personal-level truths won't allow that Jane's predicament is anything like the one that Hume puts us in. They may say that Hume follows Descartes in separating mind from brain, and they may attribute his strange claims to this. But notice that there is actually no mention of a mind in Hume. Hume is hostile to substances, whether mental or physical; and he is happy to assume that volitions set the animal spirits in motion. His expressions of mystery and unintelligibility relate not to the operation of mind but to the production of motion by events in the brain and nervous system. Hume's reason for thinking that our only power is to do things that we do not intend to do is that volitions, being at a distance from the limbs, are in no position to move the limbs directly. When causation is pictured, as it is by Hume, as proceeding always from event to event by relations of contiguity, the depths of the brain are the only place for the operation of a causal power antecedent to a limb movement. With that picture in place, the only question one can raise about causal history concerns how the limb gets to move. One loses sight of questions about the agent, and why she did what she did.

If this diagnosis is correct, then the thing that explains Hume's strange view is the absence of a human being from his account. There is nothing inside the skull—where causality, as Hume conceives it, is to be observed—which is in a position to move anything that it might have a reason to move: there is nothing for a predicate such as 'moves the arm' to apply to. Jane's difficulties about locating actions in the scientific story have a similar source, then. Among the flux of internal events, she cannot find any event to identify with *x*'s doing something, where *x* is a human being. Presumably if Jane sought the advice of a present-day naturalist, she would be told that some tract of cerebral events adds up to an event of someone's intentionally doing something. (Compare Velleman who writes: 'One is surely entitled to assume that there are mental states and events within

an agent whose causal interactions constitute his being influenced by a reason' 1994, p.124.) But we should wonder now whether Jane is not being asked to find an agent amid the workings of the brain by fabricating out of naturalistic elements something that can do duty for a Cartesian mind. Hume felt forced to deny that we have the power to move our arms. And so it seems should Jane—unless she is allowed to take a different view in order to find human beings making movements of bits of their bodies. Until she does so, agents will seem to be alienated even from the bits of their bodies that they can move. It is unthinkable that this should be our situation.

**5.** What is it for someone intentionally to raise her arm, if it is not, as Hume says, the operation of a power of hers to affect minute things inside her?

Well, when someone raises her arm intentionally, there is (arguably<sup>4</sup>) an event of her trying to raise it, and an event of its rising. For some causal theorists of action, this will seem to be the beginning of an analysis. If it is agreed that both a trying-to condition and an arm-rising condition are necessary, the next idea will be that a causal condition—saying that *x*'s arm's rising depends causally on *x*'s trying to raise it—supplies a third necessary condition of *x*'s raising her arm intentionally. And it might be suggested that these three conditions are jointly sufficient. But anyone who knows the history of this idea will be ready with counterexamples. Perhaps a neuroscientist intervenes between *x*'s trying to raise his arm and his arm's going up, so that even though there is causal dependence of the latter on the former, it was the scientist rather than *x* who raised *x*'s arm (if anyone did). An analysis of '*x* raised his arm' would need to include a condition that specified what it would be for a causal connection to be of the right kind. (It would need to find an informative way of excluding 'internal deviant causal chains'.<sup>5</sup>)

But if we resist the kind of alienation bred by Humean thinking, we shall be satisfied with something less than an analysis. The right kind of causal connection here, we can say, is the kind there is when someone's arm's going up is an exercise of her capacity to raise her arm at will. (The neuroscientist's role in the counterexample is to pre-empt the exercise of such a capacity.) So we could say that someone raises her arm intentionally if and only if (i) she tries to raise it, and (ii) she therein exercises her capacity to raise it so that (iii) her arm rises because she tries to raise it. Possession of the relevant capacity is presupposed to

-----  
 4 The claim that one tries to do what one intentionally does may be denied. But this claim introduces nothing that is especially likely to be rejected by those whose style of causal theory I dispute. Indeed the claim can be quite welcome to my opponents: making mention of an event of the agent's trying to do something provides them with an item of a sort that may seem to them to be suited to belong among bodily movements' causal antecedents as they conceive these.

5 Mele aims to provide informative sufficient conditions for an event's being an action in order to win an argument with an opponent who is an 'anticausalist teleologist' (see Mele 2000). I think that Mele's assumption that anyone who is opposed to 'anticausalism' must provide such conditions has prevented philosophers from seeing that there is a more modest causalism than that espoused by those who tell the standard story. (Mele persists with the standard causal story when he responds to Velleman 1992, in Mele forthcoming, Ch. 10. But Mele and I are in agreement (a) that there is no single state of mind corresponding to Velleman's agency *par excellence*, and (b) that some of Velleman's problems about locating agents go away when one acknowledges that an agent is a human being who acts.)

an agent's *trying* to raise her arm. The capacity is not exercised by someone whose arm makes movements against her will—as in anarchic hand syndrome<sup>6</sup>; it is thwarted when someone is impeded in raising her arm; and it is destroyed if an arm is paralysed. We can only latch onto the facts about someone who intentionally raises her arm when we allow her to be capable of raising it.

For all its circularity, this non-reductive account may be instructive. Of course its lack of analytical ambitions ensures that it cannot be of any help to anyone who had hoped to find naturalistically approved terms for describing the events that occur on an occasion when someone raises her arm intentionally. But it is genuinely a causal account, which reveals bodily agency as involving psychophysical capacities that depend on human beings' causal complexity. It acknowledges, as Hume could not, that we have the power to move our arms. And it shows immediately that there is something wrong with the standard causal story of an action—as a belief-and-desire-caused piece of behaviour.

**6.** The standard story has no dispute with my characterisation of an action as an event of someone's doing something intentionally.<sup>7</sup> And it can be agreed on all hands that a causal explanation of a certain sort can be given of why someone did something that she intentionally did. An explanation of the relevant sort shows that, in the circumstances the agent found herself or took herself to be, doing the thing was warranted or seemed to her to be. Those who are prepared to stretch the idea of 'having a reason' somewhat put this by saying that the agent had a reason to do the thing. And those who have a simplistic conception of a reason will then think of *x*'s having a reason for  $\phi$ -ing as *x*'s having a desire which *x* thinks will be satisfied if she  $\phi$ -s.

So far, there need be nothing wrong with this story, beyond its overgeneralizing the role of *reasons* and of *desires*. But it is from this story that naturalists reach 'the standard causal story'. They do so by converting the claim that *x*'s  $\phi$ -ing is explained by her having a desire which she thinks will be satisfied if she  $\phi$ -s into the claim that belief-desire pairs cause bodily movements. The conversion takes place by way of three transitions. First, *x*'s desiring something and believing something is translated into talk of items with causal potential, so that *x*'s having a reason is taken to be a matter of the existence of a pair of states.<sup>8</sup> Secondly, the fact that these states are cited in a causal explanation of why *x* did what she did is taken to be equivalent to their being causes of an action. And thirdly, an action is thought of as something on the physical side of a supposed mental/physical divide and called a bodily movement.

-----  
6 Anarchic hand syndrome is a rare condition, owed to injuries to the motor area of the brain and *corpus callosum*, from which Dr. Strangelove (the Peter Sellers' character in the Kubrick film of that name) suffered.

7 My way of telling the story assumes that Davidson is right about actions' individuation. There are naturalists who think that Davidson is wrong about that. My claims would need to be recast to count against them.

8 Of course there is a use of 'state' such that a person's desiring something or believing something (not to mention having a capacity to move her arm) is a state of hers. But in this use of 'state', states don't belong in a category of particulars which includes events—or, as Velleman says, 'occurrences, the basic elements of explanation in general .. in terms of which any explanation of human action will speak (1992, p.130). Steward 1997 contains very effective criticism both of the way that 'state' has come to be used in philosophy of mind, and of the model of causality that is brought to the subject with a 'naturalistic conception of explanation' such as Velleman's.

The central claim of the standard story of action, then, in its most familiar version, is that belief-desire pairs cause bodily movements. When ‘try to’ is introduced, the claim may be that a “belief-desire pair” causes “a trying”, which causes a movement in its turn.<sup>9</sup>

There is no need to look at further details to see that the story creates the problem which Velleman called the problem of agency. In relying on the idea of items linked in a causal chain, the standard story treats causation as Hume did, and takes it to be possible to find an action without locating any bodily being who can move. We cannot then see any agent making any difference to anything: we have the problem ‘of finding an agent at work amid the workings of the mind’.

**7.** Velleman saw no difficulties about *bodily* agency when he posed his problem of agency. That problem is supposed to arise specifically in cases where the agent does not suffer from being or feeling alienated, whereas there is bodily agency—of which Hume gives such a strange account—whether or not the agent is alienated. The standard causal story is fine with Velleman so long as it is told about agents who are depressed or fatigued, or who lack control, or would prefer to be motivated differently from how they actually are.

Inasmuch as Velleman’s own problem of agency is restricted to what he calls agency *par excellence*, Velleman must think that the picture got from Nagel’s external perspective, from which human beings are absent, succeeds in containing the truth about agency at least some of the time. But one wonders then how Velleman can think that Nagel’s external perspective reveals ‘the obstacle to reconciling our conception of agency with the possible realities [given our] scientific view of the world’ (p.129). For as Nagel saw things, there appears to be no room at all for agents in the naturalistic explanatory order.

Well, part of the explanation of Velleman’s belief that the standard story sometimes has application is his thinking that there is agent participation of a sort wherever there is human action. He says that ‘every action must be ... such than an agent participates in it, in the sense that he does it’ (n.5, p.128). Here he relies upon assuming that we encounter a trouble-free kind of agency as soon as we can say ‘He does it’. But in making this assumption, Velleman refuses to face up to the threat of unthinkable alienation. Nagel said that his external perspective presented a *general* threat to “he does it” being true in any sense; and we saw that this seems exactly right if we construe ‘he did it’ as expressing a relation between a person and an event in the naturalistic explanatory order (cp. §3). It is true that we also saw that ‘He did it’ is not actually understood in this way, and that Nagel’s threat is engendered by a misunderstanding about how people relate to the events that are actions. But the present point is that Velleman cannot consistently hold both that Nagel’s external perspective poses some genuine threat to our agency, and that someone’s action is unproblematically accommodated in the naturalistic explanatory order by virtue of his having “done it”.

In fact Velleman appears to acknowledge that there is a more general problem than the one he labels the ‘problem of agency’. For he tells us that the *mind-body problem* is that of ‘finding a mind at work amid the workings of the

-----  
 9 My own claim has been that a bit of a person’s body’s moving may be causally dependent on her trying to do something. With the notion of ‘trying to’ introduced, then, it can be all right to speak of ‘bodily movements’. But this is not to say that it is all right to speak of actions themselves as bodily movements—to speak as if a person’s doing something were a bit of her body’s moving.



body' (p.131). And if we are to think of 'a mind' as something that may move the body, then the phenomenon of bodily agency presents us with the mind-body problem as Velleman thinks of it and as we encounter it so vividly in Hume. (Notice that given this account of the mind-body problem, the problem Velleman labels the problem of agency is a problem about locating something amid the workings of something which has a problematic location amid the workings of the body.)

However exactly Velleman arrives at his view of the extent of the problem that he labels the 'problem of agency', he proposes to solve it by introducing a particular mental state, not usually admitted by the standard story's advocates, among actions' causal antecedents. The state in question is one that 'plays the functional role of an agent', and Velleman postulates that such a state is operative when there is 'agency *par excellence*'. But we can see now that this proposal is not addressed to the standard story's real difficulties. If human actions cannot be located among states and events viewed as part of 'the flux of events in nature', then introducing another state into that same flux could never be a recipe for bringing them in. A state supposedly playing the functional role of an agent brings too little too late. Such a state is literally too little, because full-sized human beings, not merely the putative inhabitants of their minds, are agents. Such a state arrives on the scene too late, because, as we shall see, human beings are thoroughly involved not only in their actions but also in their actions' causal pasts.

**8.** Velleman thinks that an agent's feeling or being alienated from what she does is a matter of the relation between her and her action falling short of what is required for a case of real agency, or of 'agency *par excellence*'. But if, as I have argued, the agent-action relation is simply that between *a* and *a*'s doing something, then it is impossible to make literal sense of this. And if, as I have argued, human beings are actually ineliminable from an account of their agency, then someone who fails to exhibit agency *par excellence* cannot be treated as someone in whom some functional role state fails to do its bit. Evidently we need to think differently from Velleman in order to draw distinctions between alienated, non-full-blooded agency and agency *par excellence*. But we shall discover that we naturally think differently: it is only to those in the grip of the naturalistic conception of what happens when someone acts that it could seem that differences between actions had always to be recorded as differences between causally efficacious items that produced them.

The agents for whom Velleman thinks the standard story is adequate are (as we saw) people who lack self-control, or self-understanding, or who fail to act wholeheartedly. In these cases explanations may appeal (respectively) to the strength of the agent's desires, to the impotence of her reasons, and to the force of an emotional reaction that she herself has not fully acknowledged. But in none of these cases should we succumb to thinking of states and events which are items inside her and which cause her body's movements. When someone's springs of action are ones she would prefer be rid of, it is understandable that we should liken them to constraints, and it is true that the language of forces and inertia then comes very naturally. But a person who appreciates that her conduct is out of accord with what she values, or is swayed by factors whose influence she regrets, admits her own motivations even if she does not approve of them. The desires and emotional states which explain what she does are after all states of *hers*—of the human being whose capacities to make movements are exercised—and, even where she feels alienated from them, they are not adventitious forces in her brain. (To think of adventitious forces in the brain seems more appropriate in

understanding, say, the involuntary movements of sufferers from anarchic hand syndrome, which lack any personal psychological explanation.)

Of course we must allow that an agent can be, or feel, more or less alienated. To allow for this, we might think of agency as coming in degrees. There is a range of properties possessed by agents which they may exhibit more or fewer of on occasion, and to a greater or lesser extent. Our conception of an agent-in-the-highest-degree might be a conception of someone who is fully self-reflective and has complete self-control, who has values and makes valuational judgements upon which she acts, who uses reason and argument effectively, who is sensitive to her circumstances, who puts her heart into what she does, and who, as we say, *identifies* with her motivations and with what she does.<sup>10</sup> To the extent to which a person's doing something on an occasion shows her as deficient compared to an agent-in-the-highest-degree, we could think of her as failing to participate in Velleman's agency *par excellence*. This is now to think of her as falling short of some ideal or other, and not as lacking some causally potent brain state.

A particular division among agents' properties will be important in considering what might be demanded of an agent in the highest degree. For we want to distinguish between the agency of mere animals, the regulation of whose lives follow biological patterns, and the agency of self-determined, human beings. In the animal case, drives, instincts and desires of certain sorts loom large in the aetiology of behaviour. In the human case, the influence of reason is characteristic. Yet even where the agent is a human being, what is done sometimes fails to be caught up with the appropriate functioning of a reasonable being and is then explicable in a more or less animal mode. Acting out of an addiction would be a case in point. We can understand why someone who has the capacities of a human agent may feel distanced from what is thus explicable, and why one should think that her agency then is less than full-blooded agency. But this is not to follow Velleman in thinking that the standard story can be told. For we still have an agent, something she does, and a psychological account of that.

Velleman sometimes writes as if his problem of agency were a problem about setting human action apart from the rest of animal behaviour. 'What makes *us* agents', he says, '—in our conception of ourselves, at least, if not in reality—is our perceived capacity to interpose ourselves into the course of events in such a way that the behavioural outcome is traceable directly to us' (p.128). Provided that 'interposing ourselves' is understood here as a matter of exercising our distinctively human capacities, and standing as we do to the events that are our actions, this seems exactly right (cp. end of §3). But in that case Velleman's qualification 'if not in reality' is surely needless. And there can be no need to deny that animals too can exercise their (animal) capacities (although the events of their exercising them are evidently not human actions).<sup>11</sup>

-----  
 10 The idea that an agent can be more or as less fully identified with what she does is one of the things that leads to Velleman's talk of the agent/action relation as coming in degrees. The language of *identification* crops up sometimes in the personal identity literature, and there it sometimes leads, as it sometimes does in the present case, to an unwarranted reductionism.

11 The picture from Nagel's external perspective in which only events and states are visible extrudes non-human animals along with human beings. For something about how human beings may be accommodated into a different picture as, as it were, a special sort of animal, see McDowell 1994, pp.114–119, and M. McGinn 2000, esp. pp.309–315.

Philosophers are interested in human beings. And they are interested inevitably in defects of agency—whether lack of self-control, weakness of will, failures of self-understanding or features that impugn responsibility. This interest encourages one to forget about the mundane and habitual. When someone, say, puts on her coat, leaves the office, and buys the evening newspaper before getting on the bus which will take her home, she does not express any deeply held values, or deliberate very much, or display particular self-knowledge or self-control. But nor is her agency defective: there are straightforward explanations of what she does, and even if these allude only to mental states of kinds recognised in the standard story, we feel no pressure to add an extra ingredient in order to reveal her as a more or less reasonable, conscious being. Velleman leaves out the relatively mundane when he contrasts various kinds of defective, alienated agency with agency *par excellence*. The omission presumably stems from his thinking that the standard causal story needs some special supplementation if it is to contain a genuine agent. But when human beings themselves are an acknowledged part of the subject matter in explanations of things they do, there ought to be no pressure to add special states of mind, beyond those that are ordinarily recognised, to ordinary explanations. We can then understand why, in unremarkable cases, a person does not have to exhibit *any* of the properties that one might associate with ‘agency *par excellence*’. And we shall find no reason to think, as Velleman does, that there has to be some single line to be drawn between defective agency and the real thing—some one state that makes human beings the sorts of agents that they are. It would actually be very remarkable if someone could exhibit *all* of the properties of an agent-in-the-highest-degree [all at once, as it were]—which is what a case of agency *par excellence* seems to demand.

One can understand Velleman’s special interest in agency *par excellence*. For we certainly don’t wish everything we do to be the product of desires we share with non-human animals, or to be a matter of habit or routine. But inasmuch as we do aspire to participate in agency *par excellence*, that need not be because we hope that some particular mental state should be operative in us as often as possible, but because we hope that we have all those standing capacities that we associate with agency in the highest degree. The extent to which we should wish actually to exercise such capacities obviously depends upon the kind of conduct that might be called for from us on occasion. An account of agency *par excellence*, then, can be focused on what should be contained in a description of a human being ideally equipped for life’s contingencies. There is no need to think of it as an account of a particular sort of psychological machinery at work on each and every occasion of action when the agent is not alienated.

**9.** There are faults in the standard causal story (as I sketched it in §6) of a sort that I have not spoken to here. Nearly everyone would agree that an adequate account of human motivation would include mental states of many more kinds than the standard story recognises, and that it is the product of overgeneralization. It is widely accepted, for instance, that people’s having intentions and plans cannot be reduced to their having reasons.<sup>12</sup> And not only (as I suggested above) does the notion of ‘a reason’ have to be stretched if human agency is always to conform to the standard story, but it is also true and widely

-----  
 12 Michael Bratman argued this, and demonstrated the shortage in the kinds of mental state that the standard story trades in, in his 1987.

acknowledged (and it leads to an opposite sort of distortion) that with its casting of *desire* as a ubiquitous motivational ingredient in the genesis of action, the standard story obliterates distinctively rational and deliberational influences on an agent's conduct. Sticking with the standard story's conceptual resources then has the consequence that someone who acts always out of non-reasoned desires can be a paradigm of a human agent. This could provide another part of the explanation why Velleman takes the story to be adequate to telling us what happens when someone acts in some defective way, but to have peculiar difficulties when it comes to agency that is distinctively human or especially full-blooded.

But once it is allowed that full-blooded human beings are the topic of an account of human agency, the project of providing an account will not seem to be that of adding further pieces of psychological machinery to states of belief and desire. An account of human agency that is allowed to be a part of an account of human beings can speak of states of mind from a broad range—virtuous or vicious traits of character, dispositions of evaluation, patterns of emotional reaction, personal loyalties, and commitments that derive from people's various individual projects.<sup>13</sup> It is easy to avoid the distortions that the standard story introduces with its supposition that everything we can know about our nature as practical beings is to be incorporated in a psychological theory that speaks of our inner workings.

**10.** Let me try to sum up, and reach the main conclusions.

I have claimed that there is alienation of an unthinkable sort when an agent is portrayed as if she were merely an arena for events. And I have also claimed that the project of looking for an agent amid the workings of a mind could never assist in getting rid of such alienation. No-one ever does anything in Nagel's picture, and it could hardly make any difference to this which particular kinds of states and events are supposed to be present from the external perspective. From that perspective, the events which are actions are missing, and they cannot be introduced by postulating a special kind of cause for them.

I suggested that there is something peculiar about thinking that the standard story of agency encounters a particular problem in cases of non-defective agency. How could it be that the story is fine so long as it is told about agents who are depressed or fatigued, or who lack control or self-determination, or who would prefer to be motivated differently from how they actually are? Our understanding of such agents relies upon our knowing that they lack some capacity, or are unable to, or fail to, exercise some capacity. But then we understand them as beings who might have possessed, or have exercised, the relevant capacities; and their status as human agents is presupposed to their conduct's being explicable as it is. Someone who falls short of displaying the properties of the paragon agent on some occasion is not treated as if they were then simply the locus of series of mere happenings.

-----  
 13 I quote more or less from Bernard Williams's description of the ingredients in a person's 'motivational set' in his 1981. Williams 1987 is a good antidote for those who are apt to think that there is some one, significant line to be drawn between agency that is genuine/full-blooded/par excellence and agency of a defective sort. But beware: Williams means something different both by 'naturalism' and by 'action' from what I use these to mean for the purposes of the present paper: (see §1 and parenthetical paragraph in §3 above).

If one ignores the gross physical facts of bodily agency, then it will be relatively easy to suppose that an agent's participation requires nothing that is obviously missing from the standard story of states and events as causes. But it becomes clear that Velleman's problem of agency would be a general problem (if it were a genuine problem at all) when one considers bodily agency, and encounters the species of unthinkable alienation introduced by Hume.

Velleman himself puts his problem of agency in a quite general way at one point: he says that it is difficult to know 'how the existence and relations of .. mental states and events .., connected to one another and to external behaviour by robust causal relations, .. can amount to a person's causing something rather than merely to something's happening in him'. To this the answer now is simple: 'They cannot'. No compounding of states and events in the naturalistic picture from which human beings are absent could constitute someone's doing something that she intentionally does.<sup>14</sup>

Velleman assumes that only *mental* states feature in the causation of actions, and he treats *states* as things to be lumped together with events in a single ontological category of 'items' or 'occurrences' (p.130). The assumption might seem to be recommended and the treatment necessary if the causal dependencies recognised when people do things had to be discernible among the causal chains that constitute the world's naturalistic workings. But the real causal dependencies are not discernible there. Nor can they be introduced by superadding a surrogate for a human being on top of—or (more literally, as Velleman sees things) in the middle of—the standard causal story.

AFTERWORD ON VELLEMAN. In his 'Introduction' to *The Possibility of Practical Reason* (Oxford University Press, 2000), Velleman no longer wants to draw exactly the line that he aimed at in the 1992 paper discussed above. And Velleman no longer speaks directly to the question whether states of agents' minds which are introduced to characterize different kinds of agency must belong within an account that subscribes to only a naturalistic conception of explanation. But some of my criticisms still apply, I think. In the 'Introduction', Velleman retains the idea that there is some *one* property of agents that we must uncover to characterise autonomous action. And he carries on with the idea of 'adding to the standard model' (pp.10–12). His thoughts about 'a mechanism modifying the motivational forces [already] at work' in a creature not endowed with practical reason also show him as captive still to the conception under attack in the present paper, I think.

The distinction that Velleman wants to capture in the 'Introduction' is between autonomous action and mere activity (as opposed to the 1992 distinction between agency *par excellence* and something relatively defective). Mere activities include so-called sub-intentional cases (along with the cases of defective agency of concern in 1992). This means that the category within which the new distinction is to be made is not the category of actions as I have characterised these here, using 'intentionally'. Still, the crucial line on which I should insist is that between cases where the agent belongs in the story and cases where she does not. Thus I would agree with Velleman that one could find fault with the standard story for its assumption that the only

-----  
 14 There is much more to be said against the psychological reductionism to which Velleman thinks we are obviously entitled: see e.g. Dupré 1993, esp. Chs. 4 and 7. For those who have joined the naturalists in their habits of thought, it might help to point out that even when one contends that there are facts which are not part of the world defined by their naturalism, plenty of materialist intuitions can be retained: see e.g. Haugeland 1984.

important line to be drawn comes between actions (as characterised here) and other events. (This doesn't come to the surface in the present paper, but it will have repercussions for how one thinks of the 'personal level' of explanation.)

Velleman has been a suitable person for me to single out for criticism because a naturalistic metaphysics informs his work even though the questions he addresses are not stock questions in theory of action. Certainly his work on agency in the last decade (see papers in 2000 collection and subsequent papers) contains much which is of immense interest and which can be disentangled from the naturalistic thinking that I have criticized here.

### References

- Bratman, Michael, 1987: *Intention, Plans and Practical Reason*, Harvard University Press, Cambridge Mass.
- Davidson, Donald, 1973: 'Freedom to Act', reprinted as Essay 3 in *Essays on Actions and Events*, Oxford University Press, Oxford, 1980, pp.63–81.
- Dupré, John, 1993: *The Disorder of Things*, Harvard University Press, Cambridge Mass.
- Haugeland, John, 1984: 'Ontological Supervenience', *Southern Journal of Philosophy* 22, 1–12. (Reprinted as Ch. 6 of *Having Thought: Essays in the Metaphysics of Mind*, Harvard University Press, Cambridge Mass., 1998.)
- Hornsby, Jennifer, forthcoming: 'Disappearing Agents', in a volume based on the 2002 Royal Institute of Philosophy Conference: *Agency and Action*, ed. Helen Steward, Cambridge University Press, Cambridge.
- Hume, David, 1748: *An Enquiry Concerning Human Understanding*.
- McDowell, John, 1994: *Mind and World*, Harvard University Press, Cambridge Mass.
- McGinn, Marie, 2000: 'Real Things and the Mind-Body Problem', *Proceedings of the Aristotelian Society* 100, pp.303–317.
- Mele, Al, 2000: 'Goal-Directed Action: Teleological Explanations, Causal Theories, and Deviance', *Philosophical Perspectives*, 14, pp.279–300.
- Mele, Al, forthcoming: *Motivation and Agency*, Oxford University Press, Oxford.
- Nagel, Thomas, 1986: *The View from Nowhere*, Oxford University Press, Oxford.
- Steward, Helen, 1997: *The Ontology of Mind: Events, States and Processes*, Oxford University Press, Oxford.
- Velleman, David, 1992: 'What Happens When Someone Acts?', *Mind* 101, pp.461–481. Reprinted at pp.123–143 in Velleman 2000 collection, to which page references here refer.
- Velleman, David, 2000: 'Introduction', Velleman 2000 collection, pp.1–31.
- Velleman, David, 2000 collection: *The Possibility of Practical Reason*, Oxford University Press, Oxford.
- Williams, Bernard, 1981: 'Internal and External Reasons', in *Moral Luck*, Cambridge University Press, Cambridge, pp.101–113.
- Williams, Bernard, 1989: 'Voluntary Acts and Responsible Agents', *Oxford Journal of Legal Studies*, reprinted as 2. in his *Making Sense of Humanity and other philosophical papers*, Cambridge University Press, Cambridge, 1995.