

# 1

## Agency, Ownership, and the Standard Theory

Markus E. Schlosser

The causal theory of action has been the standard view in the philosophy of action and mind. In the philosophy of mind, it is a piece of orthodoxy that is widely taken for granted and hardly ever questioned. In the philosophy of action, it has always had its critics. In this chapter, I will present responses to two challenges to the theory. The first says, basically, that there is no positive argument in favour of the causal theory, as the only reason that supports it consists in the apparent lack of tenable alternatives. The second challenge says that the theory fails to capture the phenomenon of agency, as it reduces activity to mere happenings (events and event-causal processes). This is often referred to as the problem of *disappearing agency*. A full defence of the causal theory should address both challenges. In the first part of this chapter, I will present what I take to be the core of the causal theory. In the second and the third part, I will then offer my responses to the two challenges. I will present a positive argument for the causal theory on the basis of considerations concerning the metaphysics of agency, and I will suggest that we *own* the agency that springs from our mental states and events *by default*.

My main aim is to show that there is no problem of disappearing agency, and we will see that my response to the first challenge will be conducive to this end. Let me point out, right at the start, that there are many controversial issues concerning the metaphysics of action and causation that are beyond the scope of this chapter. This means that I will have to make some substantial assumptions, especially in the first and second part.

### 1.1 Part 1: The standard theory

The causal theory of action, as I understand it, consists of two parts: a causal theory of the nature of actions and a causal theory of reason explanation. The former says, basically, that actions are events with a certain causal history (they are actions *in virtue of this history*). The latter is often stated by reference to Davidson's claim that reason explanation is a "species of

causal explanation" (1980, p. 3). The core of the causal theory can then be unpacked as follows:

Causal theory of the nature of action (CTA): An agent-involving event is an action just in case it is caused by the right agent-involving mental states and events in the right way.

Causal theory of reason explanation (CTR): Reason explanations of actions are explanations in terms of the agent's mental states and events that rationalize and causally explain their performance. In particular, a reason explanation of an action in terms of the agent's mental states and events is true only if those states and events causally explain the action.

The *right* mental states and events, which CTA refers to, are mental attitudes that *rationalize* the performance of the action (such as desires, beliefs and intentions). The *right way* of causation is *non-deviant* causation. We will return to the issue of non-deviant causation in the third part. For now, consider the following clarifications.

Firstly, it is plausible to think that CTA and CTR stand or fall together, because it is plausible to assume that the mental states and events that CTA refers to are the same mental states and events that feature in reason explanations according to CTR. I will assume, throughout, that CTA and CTR do stand or fall together.<sup>1</sup>

Secondly, the conjunction of CTA and CTR is only the skeleton of the causal theory. It is supposed to capture the common core of causal theories of action or, more generally, the core of the causal approach to agency. In the following, I will nevertheless refer to this core as a *theory*, for the sake of convenience.

Thirdly, the view is usually referred to as the causal theory of action. **Strictly speaking, it is an *event-causal* theory. There are alternative causal theories of agency, and I shall refer to it as the *event-causal* theory, or the *event-causal* approach, in order to avoid misunderstanding.** As I understand it, an event-causal theory is not committed to the claim that only events are causally efficacious entities. It may allow for the causal efficacy or relevance of states, dispositions and other standing conditions, and it may construe events as particulars or instantiations of properties. But it is committed to the claim that at least one event is among the causes of every effect. Events are thought to play a central role in every causal transaction, because events are the entities that initiate or trigger the occurrence of the effect. **Another way of characterizing the kind of causation in question would be to say that it is *efficient* causation by events, states or property-instantiations (by substances) as opposed to both *teleological* causation and causation by substances (*qua* substances).**

Fourthly, CTR makes no claims concerning the nature of *reasons*. In particular, it does not claim that reasons are identical with mental states or

events, and it is compatible with the view that reasons are facts or states of affairs (see Dancy 2000, for instance). CTR says that reason explanations are given in terms of mental states and events that rationalize the action. If reasons are facts, CTR can accommodate this by construing their role in reason explanations as the contents (or as what is being represented by the contents) of rationalizing mental states and events (in true and justified reason explanations). In order to avoid misunderstanding, I will call mental attitudes that rationalize actions *reason-states*, rather than reasons. Given this, any attempt to reject the event-causal theory by objecting to the claim that reasons are mental attitudes is simply missing its target.

Finally, the theory is sometimes referred to as the desire-belief theory of action. This is misleading insofar as the event-causal approach is by no means committed to the claim that all actions are caused and explained by desires and beliefs. Most proponents of the view hold now that intentions, construed as a genuine type of mental attitude, play a central role in the causation and explanation of action.<sup>2</sup> In principle, proponents of the view may refer to all kinds of mental entities that qualify as an agent's mental states, mental events, and to relationships between them. My aim here is to defend the event-causal approach to agency in general, rather than a particular version of the view.

## 1.2 Part 2: Why believe it?

Is the event-causal theory a plausible and intuitively attractive position? **It is, I think, neither particularly plausible nor particularly implausible on the grounds of commonsense intuition. The reason for this is that commonsense is silent about most of the issues involved.** It seems very plausible to say that the explanation of an action in terms of the agent's desires, beliefs and intentions is true only if the cited mental states motivated the agent to perform the action. But does this mean that the agent's mental states must have *caused* the action? Does this commit us to the claim that actions are *events*? Commonsense holds no answers, let alone conclusive reasons in favour of a particular causal or non-causal position.

What, then, is to be said in favour of the event-causal theory on the grounds of philosophical argument? Many proponents of the view think that the best argument is provided by Davidson's challenge from reason explanation. Davidson pointed out that it is possible to rationalize the performance of an action in the light of some of the agent's reason-states irrespectively of whether or not the agent acted *for* them. The agent may not have performed the action *because of having* those reason-states. In other words, reason explanation cannot be reduced to rationalization. Something else is needed.

This point is particularly salient in cases in which an agent has more than one set of reason-states that would rationalize the performance of an action,

but in which there is reason to think that the agent acted only because of *one* set of reason-states. What explains the fact that the agent acted for one rather than another set of reason-states? In general, and to use Davidson's expression, what is the "mysterious connection" between reasons and actions? Davidson suggested that this can only be a causal connection. What else could it be? This, in short, is Davidson's challenge. (Davidson 1980, especially pp. 8–11. Compare also Ginet 2001.)

It has been pointed out that the core of this challenge is a metaphysical rather than an epistemological point (Child 1994, for instance). It is not about how we can know that an agent acted for a reason, but about what makes this true. Causal connections can ground the truth of reason explanations, and they can, in general, explain the metaphysical connections between an agent's reason-states and actions.

Opponents have responded by pointing out that this challenge provides at best an indirect and negative argument for the event-causal approach. We can distinguish two points here. First, as an inference to the *only* explanation, the argument lends merely *negative* support. It supports the view only on the assumption that there are no viable alternative accounts of reason explanation that can meet Davidson's challenge. Secondly, it is only an *indirect* argument insofar as it gives no direct and positive argument for the metaphysical framework of the event-causal approach. In particular, it offers no direct support to CTA. It provides negative support to CTR, and gives indirect support to CTA only insofar as CTA and CTR stand or fall together.

Given all this, it seems that one can undermine the force of the argument simply by presenting an alternative theory of reason explanation that meets Davidson's challenge. Wilson (1989), Sehon (2000), Ginet (1990 and 2001) and Schueler (2003) have pursued this line of argument, and they have offered alternative non-causal accounts of reason explanation. Proponents of the event-causal theory have responded by criticizing the offered non-causal alternatives, and they have argued, convincingly I think, that Davidson's challenge is very much alive (compare Mele 1992 and Clark 2003, for instance). But I do not think that the case for the event-causal approach rests on Davidson's challenge alone. In the remainder of this part, I will present a direct argument from the metaphysics of agency.

### 1.2.1 Naturalism and the event-causal order

Many philosophers, I suspect, would not agree with the suggestion that Davidson's challenge provides the only argument in support of the event-causal theory. For many, the best reason to endorse the view consists in its apparent compatibility and congeniality with naturalism. According to a fairly widespread form of philosophical naturalism, all particular occurrences, processes, and changes are to be understood and explained in terms of event-causation. In particular, any appeal to substance-causation, irreducible teleology or *sui generis* acts would constitute a violation of naturalism.

Given a strong commitment to this kind of naturalism, one has, it seems, a strong reason to endorse the event-causal approach to agency, because this view *locates* or *situates* agency within the event-causal order, as it were. Opponents of the view will point out, rightly, that this does not give them any reason to endorse the event-causal theory, simply because they reject this kind of naturalism. (This does not commit them to some kind of anti-naturalism. The term *naturalism* has been used in many different ways. Rejecting the outlined kind of event-causal naturalism, opponents of the event-causal theory need not reject naturalism as such.)

However, if the characterization of naturalism is weakened, in the right way, we can construct an argument for the event-causal theory that does not beg the question. Or so I will argue now. It should, I think, be uncontroversial that human agents are part of the event-causal order in the following weak sense. Our bodily movements are events that are part of the event-causal order in the sense that their occurrence and execution can be explained in terms of event-causation only—in terms of muscle contractions, motor commands, neural activity in the motor cortex and areas of the prefrontal cortex, and so on. This claim is clearly weaker than the claim that all occurrences, including actions, must be explained in terms of event-causation, and I suspect that only very few non-causalists, if any, would object to this weak or minimal version of naturalism. It is undeniable, I think, that we (or our *living bodies*, if you like) are part of the event-causal order in this weak or minimal sense. I will assume, from now on, that this is the case, and I will express this by saying that human agency is *minimally part* of the event-causal order. This is the first of two main assumptions for my argument.

Non-causalists often point out that we are primarily interested in actions, rather than bodily movements, when we are interested in human agency. When we give a reason explanation, for instance, we explain the performance of an action, rather than the mere occurrence of a movement. This, they rightly point out, is often overlooked due to the close relationship between bodily movements and basic actions. Roughly, a *basic action* is something an agent can do without doing anything else.<sup>3</sup> It is widely agreed that every action of ours is either a basic action or brought about by the performance of a basic action (perhaps via a chain or tree of non-basic actions). To take a standard example, you can give a signal by raising your arm. If you do so, you perform a non-basic action (giving a signal) by performing a basic action (raising an arm). The basic action is not performed by doing something else. It is, in particular, not performed by performing a bodily movement (the rising of your arm). But the basic action is also not identical with the bodily movement. Not every movement of this type constitutes or realizes the raising of an arm. They may be token-identical, but they are not type-identical.

Proponents of the event-causal theory can agree with all this, and I think they should agree. We are interested primarily in intentional behaviour, not

bodily movement. We perform all non-basic actions by performing basic actions, and basic actions are not type-identical with bodily movements. Nevertheless, it seems obvious that basic actions stand in some intimate relationship with bodily movements. Should they not be token-identical, basic actions are at the very least partly constituted or realized by bodily movements.<sup>4</sup> This appears to be obvious, and I will assume that this is the case. That gives us the second assumption for my argument.

Taken together, the two assumptions give rise to the following central question in the metaphysics of agency: how can human agency be part of the event-causal order at all? Given that overt actions are constituted or realized by movements, and given that bodily movements can be explained in terms of neuro-physiological events, how can agents exercise their agency? How can agents, who are minimally part of the event-causal order, exercise control over their actions?

To make this question clearer, let us briefly consider the case of actions that are done for reasons (in the broad sense of being rationalized and motivated by reason-states). We can distinguish here between four things: actions, an agent's reason-states, bodily movements, and the causes of bodily movements. On the one hand, these actions are done for reasons. They are performed because their agents have certain reason-states. On the other hand, they are at least partly constituted or realized by bodily movements, which can be explained by reference to neuro-physiological events alone. Given the constitutive relationship between actions and movements, we need an explanation of how the influence of reason-states on the agent's actions is related to the causal efficacy of the neuro-physiological causes of the agent's bodily movements.

Agential and rational control is in need of explanation. In the following section, I will introduce what I take to be the basic options in the metaphysics of action, and we will assess them in light of the task that has just been outlined.

### 1.2.2 The metaphysics of agency

What is an agent? What is agency? First, let me restrict our considerations to overt actions (that is, roughly, actions that involve bodily movement).<sup>5</sup>

A good starting point is to think of overt agency in terms of self-movement. Agents are beings or systems that can bring about change in their environment by bringing about change in themselves (by moving in a certain way). Agency is an exercise of this ability.

What is self-movement? Intuitively, it is movement that is brought about or initiated by oneself (by the agent or the system itself). What does this mean? We can distinguish here between three main options in the metaphysics of agency, which give three different answers to this question.

According to the first, self-movement is initiated by the self in the sense that it is caused by salient features of the agent, which are themselves caused

Agents  
and  
agency

3 options  
for agency

by other features of the agent or the agent's environment. According to the second, self-movement is initiated, literally, by the agent (by the persisting being that is the agent, rather than features of that being). According to the third option, self-movement is initiated by a mental act of the will, often called *volition*, which is not itself caused by anything else.

Those, I think, are the three main options in the metaphysics of agency. The first is a reductive approach. It reduces an agent's role or power to the causal roles or powers of agent-involving states and events. Paradigm instances of this approach are event-causal theories, which provide a metaphysical account of agency in terms of event-causal relations between reason-states and actions.<sup>6</sup> The second approach is a non-reductive approach. It construes an agent's role or power as irreducible. Paradigm instances of this approach are agent-causal theories of action.<sup>7</sup> The third approach is also non-reductive in the sense that it rejects the reduction proposed by the first approach. Paradigm instances of this approach are volitional theories of action.<sup>8</sup>

A fourth possibility is to reject the project of giving a metaphysical account of agency as misguided. On this view, the notion of agency is essentially normative, and it can be captured and understood only from a normative or practical standpoint. The phenomenon of human agency is thought to disappear from any metaphysical, naturalistic or otherwise theoretical point of view. Let us call this the dual standpoint view.<sup>9</sup>

### 1.2.3 Dual standpoint theories

Our question is how human agency can be minimally part of the event-causal order. In particular, how can agents, who are minimally part of this order, exercise control over their actions? In order to give an answer to this question, one must provide a metaphysical account of human agency. The dual standpoint view does not acknowledge that there is a need for an explanation, and it rejects the metaphysical quest as misguided. It is, for this very reason, unsatisfactory, as it leaves one of the fundamental questions about human agency unanswered.

This dismissal of the dual standpoint view would be unjustified, if it could be shown that a metaphysics of agency *cannot* be given. But I have not seen any good argument to this conclusion. On the face of it, a metaphysics of agency is worth wanting, and different proposals are on offer. Given this, I see no reason to adopt the dual standpoint view. (We will return to the charge that agency disappears from a naturalistic standpoint further below.)

### 1.2.4 Volitionism

Volitionism is widely rejected, mostly due to internal problems with the theory. I will not summarize the most common objections here.<sup>10</sup> Let us consider, instead, whether the view can give an answer to our question.

According to volitionism, all actions originate from uncaused mental acts (acts of the will or volitions). No one, I assume, would wish to deny that mental acts are realized, partly at least, by events in the brain. But for all we know, there are no uncaused neural events. This alone, I think, renders the position very problematic and unpromising. Given that there are no uncaused neural events, neither at macro nor at micro levels of description, how could uncaused mental acts possibly be realized by neural events that are caused by other neural or physiological events?<sup>11</sup>

Furthermore, it remains a mystery just how the view can account for an agent's control at all. Volitions are neither caused by the agent, nor by the agent's reason-states. They are uncaused acts, and the agent appears to be a mere subject or bearer of volitions. Proponents of the view might say that this misrepresents their position, as volitions are necessarily *willed*, but not caused, *by the agent*—spontaneously, intentionally, and freely. But to say this is just to reaffirm the assumption that agents have control. What is needed, however, is an *explanation* of how agents can have control over actions that are constituted or realized by movements that have event-causal explanations.

### 1.2.5 The agent-causal approach

Some libertarians have tried to revive the agent-causal approach in the more recent debate on free will, because they think that only an agent-causal theory can account for the kind of power or control that is required for free will. But the agent-causal approach is still widely rejected, mainly because it is based on the very controversial notion of substance-causation (compare Clarke 2003). It has also been argued that the view is untenable in the light of empirical considerations (Pereboom 2001, for instance). In my opinion, the agent-causal approach fails more fundamentally as a theory of agency, because it fails to account for agential control. I have argued for this elsewhere (Schlosser 2008). In a nutshell, my argument goes as follows.

Both the event-causal and the agent-causal theory seek to explain agency and control in terms of causation. Of course, not every instantiation of a causal relation constitutes an exercise of control (or the instantiation of a relation of control). If a certain causal relation constitutes control, it must be in virtue of some further fact. According to the event-causal theory, control consists in non-deviant causation by reason-states. Crucial to this account are the causal roles of mental states and events, and, in particular, the causal and explanatory roles of their intentional contents. Nothing plays a similar role in the agent-causal theory. Nothing can possibly play this role, because the theory refers to agents *qua* substances as the causes of actions. On the agent-causal view, the causation of actions is not guided by any properties of the agent. But because of this, it remains obscure why we should think that instantiations of the agent-causal relation constitute exercises of control at all. Like volitionism, it does not explain agential control,



but merely reaffirms the assumption that agents have control. **Note that, according to this objection, the problem with the agent-causal theory is not that it construes agents as substances. The problem, rather, is that it does not explain how agents exercise control over their actions.**<sup>12</sup>

### 1.2.6 The event-causal approach

Proponents of the event-causal theory may identify the agent with a substance (with the living organism, for instance). But within that framework, **the agent's role, or the agent's control, is reduced to the causal roles of agent-involving mental states and events. Agential control is construed as non-deviant causation by reason-states. And the relationships between actions, bodily movements, reason-states and neural events are usually construed as intra-level and inter-level relations within a model of levels of explanation: the causal relationships at the level of mental description between reason-states and actions depend on and are realized by the causal relationships between the neural events and bodily movements that are identified at lower levels of description. In this way, the theory not only locates or situates agency within the event-causal order, but it gives an account of control and agency in terms of non-deviant causation by mental states and events. This shows how agents, who are minimally part of the event-causal order, can have and exercise control.**

This account, of course, gives rise to various questions and problems. The most pressing are the problem of deviant causal chains and the problem of mental causation (in particular, the problem of causal exclusion).<sup>13</sup> I believe that the former problem has a solution, and we will return to this below. Various solutions to the problem of mental causation have been proposed.<sup>14</sup> **Let me point out here only that it is not inherently a problem of mental causation. It is, more generally, a problem that concerns the relationships between causation and causal explanation at different levels of explanation. In particular, it is not a problem that arises especially for the event-causal theory of action. Unlike the problems for the other options in the metaphysics of agency, it is a problem that arises for all theories that assume the causal relevance of higher-level entities (including all the entities that are stipulated by the special sciences).**

This completes my case in support of the event-causal approach to agency.<sup>15</sup> **There is reason to endorse the metaphysical framework of the event-causal theory, because it is the only theory that can explain how human agency can be minimally part of the event-causal order.** It might be objected that this is also merely a negative argument, as it promotes the view only by arguing that there are no viable alternatives. But we do not have to see it that way. There is, I submit, reason to endorse the event-causal theory because it gives a *good* explanation, and this explanation also happens to be the *only* explanation. More importantly, there is a significant difference to the way in which Davidson's challenge is merely negative.

My argument provides reason to adopt the metaphysical framework that is implicit in CTR and CTA, whereas Davidson's challenge lends direct support only to CTR.

### 1.3 Part 3: Disappearing agency and disappearing agents

In this final part, I will offer a response to the challenge of disappearing agency. My response comes in two parts, which correspond to the following two versions of the challenge. According to the first, the event-causal theory altogether fails to capture the phenomenon of agency, as it reduces activity to mere happenings. Understood in this way, it is a fundamental challenge to the event-causal approach as such. Statements of this first challenge can be found in Melden (1961) and Nagel (1986), for instance. A weaker objection has been raised in the more recent debate. According to this second version, the standard event-causal theory fails to capture important aspects of human agency, because it fails to account for the proper role of the human agent in the performance or exercise of certain kinds of agency. This challenge grants that the event-causal theory can account for some basic kinds of human agency (and animal behaviour). But it calls for a substantial revision or supplementation of the view in order to account for the more refined or higher kinds of human agency. In this version, the challenge has been acknowledged even by many proponents of the event-causal approach, including Velleman (1992), Bratman (2001), Enç (2003), and Schroeter (2004). In the following, I will refer to the first challenge as the challenge of disappearing *agency*, indicating that it is a fundamental challenge to the event-causal approach to agency. And I will refer to the second challenge as the challenge of disappearing *agents*, indicating that it is a challenge concerning the role of agents in the performance of actions.

#### 1.3.1 Disappearing agency

Both challenges have been presented by means of spurious metaphors and rhetoric. According to the event-causal theory, it has been claimed, the agent is a *mere locus* in which events take place, a *mere bystander* or *victim* of causal pushes and pulls. Proponents of the fundamental challenge have sometimes used such metaphors in order to make the point that agency *disappears* within an ontology of events and event-causation.

One can acknowledge that this challenge has some intuitive force. But it is more important to note that its proponents have not produced a single argument to support their case, and they have certainly not identified a philosophical problem.<sup>16</sup> Their case is entirely based on intuition, and in some cases on mere metaphor and rhetoric.

However, having said this, and having acknowledged that the objection has some intuitive force, proponents of the event-causal theory should also be able to say something in response. It is not obvious that agency cannot

be understood in terms of event-causal processes. But it is also not obvious that agency *can* be understood in terms of event-causation. What can we say in response?

A first thing to point out is that some of the rhetoric is not just misleading, but false. The agent is certainly not a *victim* or a *helpless bystander* only in virtue of being a subject of events (in virtue of being a substance that is involved in events). Events may be called *happenings* in virtue of the fact that they *occur* in time. **But the fact that events are occurrences does not entail or show that an agent's mental events and movements are things that *happen to* the agent, in the sense that they assail or befall the agent, or in the sense in which we say that a bad or unjust thing happened to us.** When I remember something, for instance, I am a constitutive part of an event, but I am no victim or helpless bystander.

Secondly, we must remember that the event-causal theory is *intended* to be a reductive theory. Its proponents aim to provide necessary and sufficient conditions for agency, and they propose to do this without any kind of circularity by way of providing a reductive *explanation* of agency in terms of event-causal processes. But, as every proponent of a reductive explanation would insist, a reduction is a form of vindication rather than elimination. **The theory does not eliminate agency, nor does it eliminate agents. Rather, it provides a vindication by giving an account of how agency can be minimally part of the event-causal order.**

Thirdly, and most importantly, there is a constructive response to the challenge. **The challenge says, basically, that the event-causal approach fails to capture agency. We can interpret this as saying that it fails to capture the fact that agents can exercise control over their behaviour. Construed in this way, proponents of the event-causal theory can respond by showing that the view has the resources to distinguish between event-causal processes that constitute agential control and event-causal processes that do not. If this can be achieved within the event-causal framework, then the challenge is mistaken.**

### 1.3.2 Event-causation and agential control

The event-causal theory construes control in terms of event-causation and rationalization. It says that an agent exercises control only if the behaviour in question is caused by mental states and events that rationalize its performance (we call such rationalizing attitudes *reason-states*). But causation by reason-states is not sufficient for control and agency. This is highlighted by examples involving deviant or wayward causal chains.

In all standard examples of causal deviance, the causal chain that connects the agent's reason-states and the action runs through some state or event that undermines the agent's control. Typically, this is some state of nervousness or agitation.<sup>17</sup> Consider, for instance, Davidson's climber example.

A climber might want to rid himself of the weight and danger of holding another man on a rope, and he might know that by loosening his hold on the rope he could rid himself of the weight and danger. This belief and want might so unnerve him as to cause him to loosen his hold [...].

(Davidson 1980, p. 79)

Examples of this kind raise a problem for the event-causal theory. The behaviour is caused and rationalized by mental states, but it seems clear that the agent is not performing an action at all. It is, rather, a sheer accident that the state of nervousness causes precisely that type of movement that is rationalized by the reason-states. In order to provide a satisfactory account of agency, the theory must exclude deviant causal chains in event-causal terms (in particular, without presupposing an irreducible notion of control or agency).

The interesting point for us here is that deviant causal chains are *control-undermining* chains. If the theory can exclude deviant causal chains, it can, *ipso facto*, exclude control-undermining chains. And if it can exclude control-undermining chains, it can distinguish between event-causal chains that constitute agential control and ones that do not.

I think that the problem of deviant causal chains can be solved, and I have proposed a solution elsewhere (Schlosser 2007b). In broad outline, I have argued that deviant causal chains are excluded if the theory requires that the agent's reason-states must be causally efficacious and explanatory *in virtue of their intentional contents*. This requirement is violated in the standard cases of causal deviance. Given this, we get a straightforward response to the challenge of disappearing agency. The event-causal theory can capture agency, because it has the resources to distinguish between event-causal chains that constitute agential control and ones that do not.

### 1.3.3 Ownership of agency

In response, opponents might argue that the problem is not that causal chains can be deviant, but that the constituents of those chains are mere states and events. Let us assume, for the sake of argument, that actions are non-deviantly caused by the agent's mental states and events. The objection says that this still does not guarantee agency, because the agent may not identify with being moved by those states and events. Being non-deviantly caused by mental states and events, the resulting behaviour may still not be a true and proper expression of the agent's own agency. Let us call this the challenge from ownership (ownership of agency).

Many will associate this challenge with the issues raised by Harry Frankfurt's influential article on free will and personhood (1988, chap. 2, originally published in 1971). Following Frankfurt, one might be tempted to give a response to the objection by appealing to a notion of identification or endorsement within a so-called hierarchical theory of agency. This

route has been taken by some philosophers in response to the problem of disappearing *agents*. The central idea here is to distinguish between desires or motives that *speak for* or *stand for* the agent by means of an account of identification or endorsement (compare Korsgaard 1996 and Bratman 2001, for instance). On this approach, an agent acts from desires that are truly his or her *own* just in case the agent *endorses* the desires as motives for action. Let us call this the *endorsement strategy*.

I shall not attempt here to assess the endorsement strategy as such. Nor will I go into the details of particular versions of this approach. I want to argue, rather, that the endorsement strategy would not give us a convincing response to the objection from ownership.

To begin with, let us consider Frankfurt's main example of the unwilling addict. He is moved by a desire to acquire and take drugs, but he does not want to be motivated in this way—he does not endorse being moved by those desires. We can agree with Frankfurt and his followers that cases of this kind highlight important and interesting aspects of human agency. But we must be careful to interpret them in the right way.

Frankfurt says that the unwilling addict is a “passive bystander to the forces that move him” (ibid. p. 22). Elsewhere he talks about desires that are “rejected” as “external” (1988, chap. 5, for instance). David Velleman and Michael Bratman have suggested that examples of this kind show that the event-causal theory “leaves out” the agent. The unwilling addict performs an action when he acts on the desire, but this falls short of “agency *par excellence*” (Velleman 1992) or “full-blooded” human agency (Bratman 2001). Construed in this way, the example can be used to raise the challenge of disappearing agents (to which we will turn below). But I think that this reading can and should be resisted here.

Firstly, it should be uncontroversial that the unwilling addict is not a mere bystander or locus in the flow of events. He is capable of a good degree of control and agency, and he exercises this ability in the pursuit of drugs, an endeavour which requires some planning and practical reason.

Secondly, it should also be uncontroversial that the addict's desire and the resulting behaviour is his *own* in some basic or minimal sense. In order to see this, compare the unwilling addict with serious cases of schizophrenia, where patients report that their actions are under the control of some external agent or force, or with cases of the “anarchic hand syndrome,” where patients report that one of their hands moves on its own (compare Frith et al. 2000, for instance).

Third, proponents of the event-causal theory should seek to respond to the challenge at the most fundamental level. The endorsement strategy can be pursued in order to account for a kind of ownership that is characteristic of autonomous agency, for instance. But this would leave more basic kinds of agency unaccounted for. We can agree that the unwilling addict falls short of autonomous agency (or agency *par excellence*). But, on the other hand, he

is not like the schizophrenic patient who feels as if alien forces are acting through him. What is required in order to meet the objection at the fundamental level is an account of the basic kind of ownership and agency that is exhibited even by the unwilling addict.

I propose the following response to the objection from ownership. In normal instances of human agency, including basic cases of minimally rational planning agency, actions that are non-deviantly caused by the agent's mental states and events are an expression of the agent's own agency by default: our agency springs from our mental states and events, unless defeating conditions obtain. Ownership of agency, in other words, does not have to be conferred by endorsement and it does not depend on it. It is a given, unless things go wrong.

But what are normal instances? In order to get a viable response, we must have an answer to this question. Fortunately, there is a computational model of the sense of agency that provides a good answer. In broad outline, the model is this. Whenever a motor command for the performance of a bodily movement is sent from premotor areas to the motor control system, a copy of this command is used to produce a prediction of the movement (a so-called *forward model*). This prediction is then sent to a comparator where it is compared with incoming visual and proprioceptive information concerning the actual movement. The main purpose of this sub-personal system is to monitor, correct, and fine-tune movements. But it is now widely assumed that this system is also responsible for a sense of the ownership of agency. This is the sense that the movements are our own doing, initiated and guided by us, and it is assumed that this sense or feeling is the result of a match between the prediction and the feedback (the match, of course, need not be perfect, as the function of the system is to correct and fine-tune).<sup>18</sup>

There is good empirical support for this model, and it is now widely deployed by psychologists and cognitive scientists working on human action. Given this model, we can say what normal instances are. They are cases in which the feedback-comparison system performs its function, producing a sense of the ownership of agency. This sub-personal mechanism may fail to produce a sense of agency for various reasons. It may be interfered with or break down in various ways that correspond to a variety of abnormal cases and defeating conditions. What the defeating conditions are is largely an empirical question. It has been argued that the model can explain a wide range of deficiencies and abnormalities, each highlighting ways in which the mechanism may break down or fail to perform its function (Frith et al. 2000).

It would be implausible, I think, to suggest that the ownership of agency is in all cases conferred by the agent's endorsement. The correct reading of Frankfurt's unwilling addict supports this. There is a basic or minimal sense in which the addict's desire for the drug is his own and in which his own agency springs from it (in combination with other mental states and events).

computational  
model of the  
sense of  
agency

normal  
cases of  
agency

It should be noted that the proposed default view is fully compatible with an endorsement theory, as I have argued only that the endorsement strategy should not be deployed in response to the challenge from ownership. So, the default view may well be supplemented with an endorsement theory of autonomous agency, for instance.

### 1.3.4 Disappearing agents

Let us now turn to the challenge of disappearing *agents*, and let us assume that the outlined feedback-comparator model can account for a basic and default sense of the ownership of agency. What about the more refined and higher kinds of human agency, such as autonomous agency? Even *proponents* of the event-causal theory have conceded that the view fails to account for the agent's participation or proper role in the performance of higher kinds of agency.

I accept the point that the basic version of the event-causal theory has to be refined or supplemented in order to account for higher kinds of agency. But I do not accept all the claims made and implied by the challenge. In particular, I do not accept the suggestion that the event-causal theory fails to include the *agent* or fails to account for the *agent's* role and participation—a point that has been conceded even by some proponents of the event-causal approach.

Firstly, putting things in terms of the *agent's* role or participation creates a false dichotomy. Throughout this chapter, we distinguished between more *basic* and *higher* kinds of agency. It is very plausible to think that the aspects or kinds of human agency form a spectrum, or a hierarchy, from lower and basic to higher and more refined kinds of agency. At the bottom of this hierarchy one finds behaviour that is purposeful but to a high degree driven by environmental stimuli (such as instinctive, automatic or highly habitual reactions). Moving up the hierarchy we get intentional, rational, deliberative, reflective and self-controlled agency, and towards the top we find autonomous and free agency. For our purposes, the details and the exact order do not matter. The important point is that agency comes in shades of grey, as it were, not as an all-or-nothing phenomenon. Whenever human agency is exercised, some but not necessarily all kinds of human agency are instantiated.

This is why I find talk about the *agent's* role or participation unhelpful. If we say that the agent's participation is characteristic of autonomous agency, does that mean that the agent does not participate in lower kinds of behaviour? This would be rather odd, to say the least. If we want to capture the important aspects of human agency, we better begin with a framework of *kinds of agency*. Talk about the agent's role and the agent's participation creates a bipartition that does not match up with the varieties of human behaviour.

Secondly, as just pointed out, it is rather implausible to suggest that the agent does not participate in lower kinds of agency. The most natural thing

to say, and the most natural assumption to begin with, is that all instances of agency involve an agent. Wherever there is agency, there is an agent participating, playing a role as the agent. As explained in part two, the event-causal theory provides a reductive explanation of the agent's role. It does not eliminate the agent. Given that, there is simply no room for an additional role of the agent in higher kinds of agency. The agent is already there, from the start, and the agent does play a role in all kinds of agency.

In other words, to ask for further participation of the agent is to miss the point of the reductive approach to agency. Higher kinds of agency do not spring from the participation of the agent. They spring, rather, from certain features of the agent. They spring from properties that are instantiated only in cases of autonomous agency, for instance. We may say that the agent participates more, or to a higher degree, in some instances of agency. But this is metaphorical. It should be taken to mean that the agent instantiates certain properties or exercises certain abilities, which are not instantiated or exercised in lower kinds of agency, and in virtue of which the agent exercises the higher kind of agency in question.

What should we make, then, of the challenge of disappearing agents? Is it an empty challenge? Construed, literally, as a challenge of disappearing agents it is an empty challenge, as I have just suggested. But that does not mean that it is empty altogether. I acknowledge that the event-causal theory must be supplemented and refined. But not by bringing the agent back into the picture. The agent was never absent. The right way to respond, rather, is to show how the theory can distinguish between the various kinds of agency within the event-causal framework (without, in particular, presupposing some kind of agent-causation). This task is beyond the scope of this chapter. But I shall briefly indicate two directions that proponents of the causal theory may take. One possible starting point for an account of higher kinds of agency is Frankfurt's (1988) hierarchical model. In order to solve the well-known regress problems that plague this approach, one may appeal to special types of mental attitudes, such as the motive to be governed by reasons (Velleman 1992) or higher-order policies that provide cross-temporal continuity and stability (Bratman 2001). Alternatively, one can appeal to historical conditions on the way in which agency-relevant attitudes, such as desires, beliefs and intentions, must have been formed or acquired (Mele 1995). I tend to favour this second approach, as I think that higher-order attitudes play a less significant role in human agency than Frankfurt and his followers assume.

More recently, François Schroeter (2004) argued that we must refer to the role of the *conscious self* in the initiation and guidance of autonomous action. In my view, this is not an option for the committed proponent of the event-causal approach. Schroeter insists that reference to the role of the conscious self is *not* a covert evocation of some kind of agent-causation, and he claims that the view is consistent with naturalism (p. 650). I understand

adoption  
of some  
form of  
mesh-  
theory



Velleman's and Bratman's views, for instance, which account for the agent's role by reference to some of the agent's mental states. According to Schroeter, however, the conscious self cannot be reduced to conscious mental states and events, nor does he want to say that the self must therefore be a substance. But it seems clear that the self must be some kind of entity, in a metaphysically robust sense, as it is supposed to be causally relevant in the initiation and guidance of action. Schroeter does not say what kind of thing it is, and I fail to see what it could possibly be.

Perhaps the role of the conscious self goes beyond the role of conscious mental states and events due to the *unity* of the self. The question of what this unity might consist in is, of course, also beyond the scope of this chapter. I should point out, though, that proponents of the event-causal approach are not restricted to explanations in terms of collections of mental states and states. They may, rather, refer to the agent's mental states, mental events, and the relationships that hold between them. Given this, it is, I think, far from obvious that the role of a conscious and unified self cannot be captured and reductively explained by an event-causal theory of agency.

## Acknowledgements

I would like to thank John Bishop for very helpful comments on an earlier version of this chapter, which was written when I was a post-doc fellow at the University of Bristol with a research project on the causal theory of action. This project was funded by the Austrian Science Fund (FWF).

## Notes

1. This claim will appear to be rather weak if one thinks that CTR and CTA are essentially or necessarily connected. But it seems that one can coherently hold one without the other. One may, for instance, hold CTR in conjunction with a substance-causal view on the nature of actions.
2. This is usually credited to Bratman 1987. Compare also Brand 1984, Bishop 1989, Mele 1992, and Enç 2003.
3. For a recent discussion of the notion of basic action see Enç 2003.
4. I am restricting my considerations here to overt behaviour. But I think that similar claims hold with respect to the relationship between basic mental acts and neural events.
5. For an application of the event-causal approach to mental action see, for instance, Mele 1997.
6. Davidson 1980, Brand 1984, Bishop 1989, Mele 1992, and Enç 2003, for instance.
7. Chisholm 1964 and O'Connor 2000. Compare also Clarke 2003.
8. Ginet 1990 and McCann 1998, for instance.
9. Proponents include Nagel 1986, Korsgaard 1996, and Bilgrami 2006.
10. See, for instance, Brand 1984, Enç 2003, and Clarke 2003.
11. Jonathan Lowe (1993) argued that the causal relevance of volitions is compatible with the causal closure of the physical, if volitions are construed as *enabling* or

- structuring* causes rather than triggering causes. But this fails to capture the proper role of volitions. They are supposed to *initiate* actions. To construe them as enabling or structuring causes does not account for their efficacy as volitions.
12. Strictly speaking, the agent-causal approach is a *version* of the non-reductive approach to agency. I have argued that the case against the agent-causal approach generalizes to a case against the non-reductive approach (Schlosser 2008).
  13. Roughly, the problem of causal exclusion is that the causal sufficiency of physical events (or, more generally, the causal closure of the physical) appears to exclude the causal relevance of mental events (see Crane 1995, for instance). This problem arises only if one assumes that mental causation requires the downward causation of physical events. Elsewhere I have argued that the mental causation of actions does not require downward causation (Schlosser 2009).
  14. For instance Yablo 1997 and Gibbons 2006, among many others.
  15. An earlier but more detailed version of this argument can be found in Schlosser 2007a.
  16. In contrast to that, the challenge from deviant causal chains *does* raise a genuine problem. Virtually all proponents of the event-causal approach have acknowledged this. Compare Davidson 1980, Bishop 1989, and Enç 2003, for instance.
  17. I am restricting my considerations here to the most problematic type of causal deviance, which has been called *basic* or *primary* deviance. Compare Bishop 1989 and Schlosser 2007b.
  18. This is the most basic version of the model. For more advanced and more detailed accounts see Frith et al. 2000 and Pacherie 2007, for instance. For an application of this model to the case of mental agency see Campbell 1999.

## References

- Bilgrami, A. (2006), *Self-Knowledge and Resentment*. Cambridge: Harvard University Press.
- Bishop, J. (1989), *Natural Agency: An Essay on the Causal Theory of Action*. Cambridge: Cambridge University Press.
- Brand, M. (1984), *Intending and Acting*. Cambridge: MIT Press.
- Bratman, M. E. (1987), *Intention, Plans, and Practical Reason*. Cambridge: Harvard University Press.
- . (2001), "Two Problems about Human Agency," *Proceedings of the Aristotelian Society* 2000–2001: 309–26.
- Campbell, J. (1999), "Schizophrenia, the Space of Reasons, and Thinking as a Motor process," *The Monist* 82: 609–25.
- Child, W. (1994), *Causality, Interpretation and the Mind*. Oxford: Oxford University Press.
- Chisholm, R. (1964), "Human Freedom and the Self," Reprinted in G. Watson (2003) (ed.), *Free Will*. Oxford: Oxford University Press, 26–37.
- Clarke, R. (2003), *Libertarian Accounts of Free Will*. Oxford: Oxford University Press.
- Crane, T. (1995), "The Mental Causation Debate," *Proceedings of the Aristotelian Society*, supplementary volume 69: 211–36.
- Davidson, D. (1980), *Essays on Action and Events*. Oxford: Oxford University Press.
- Dancy, J. (2000), *Practical Reality*, Oxford: Oxford University Press.
- Enç, B. (2003), *How We Act: Causes, Reasons, and Intentions*. Oxford: Oxford University Press.

- Frankfurt, H. (1988), *The Importance of What We Care About*. Cambridge: Cambridge University Press.
- Frith, C. D., Blakemore, S. and Wolpert, D. M. (2000), "Abnormalities in the Awareness and Control of Action," *Philosophical Transactions of the Royal Society* 355: 1771–88.
- Gibbons, J. (2006), "Mental Causation without Downward Causation," *Philosophical Review* 115: 79–103.
- Ginet, C. (1990), *On Action*. Cambridge: Cambridge University Press.
- . (2001), "Reasons Explanations of Action: Causalist versus Noncausalist Accounts," in R. Kane (ed.), *Oxford Handbook on Free Will*. Oxford: Oxford University Press, 386–405.
- Korsgaard, Ch. (1996), *The Sources of Normativity*. Cambridge: Cambridge University Press.
- Lowe, E. J. (1993), "The Causal Autonomy of the Mental," *Mind* 102: 629–44.
- McCann, H. (1998), *The Works of Agency*. Ithaca: Cornell University Press.
- Melden, A. I. (1961), *Free Action*. London: Routledge and Kegan Paul.
- Mele, A. R. (1992), *Springs of Action*. Oxford: Oxford University Press.
- . (1995), *Autonomous Agents*. Oxford: Oxford University Press.
- . (1997), "Agency and Mental Action," *Philosophical Perspectives* 11: 231–49.
- Nagel, T. (1986), *The View from Nowhere*. New York: Oxford University Press.
- O'Connor, T. (2000), *Persons and Causes: The Metaphysics of Free Will*. Oxford: Oxford University Press.
- Pacherie, E. (2007), "The Sense of Control and the Sense of Agency," *Psyche* 13: 1–30.
- Pereboom, D. (2001), *Living Without Free Will*. Cambridge: Cambridge University Press.
- Schlosser, M. E. (2007a), *The Metaphysics of Agency* (Doctoral dissertation, University of St. Andrews). *Research Repository* (<http://hdl.handle.net/10023/163>).
- . (2007b), "Basic Deviance Reconsidered," *Analysis* 67: 186–94.
- . (2008), "Agent-Causation and Agential Control," *Philosophical Explorations* 11: 3–21.
- . (2009), "Non-Reductive Physicalism, Mental Causation and the Nature of Actions," in H. Leitgeb and A. Hieke (eds), *Reduction: Between the Mind and the Brain*. Frankfurt: Ontos, pp. 73–89.
- Schueler, G. F. (2003), *Reasons and Purposes: Human Rationality and the Teleological Explanation of Action*. Oxford: Oxford University Press.
- Schroeter, F. (2004), "Endorsement and Autonomous Agency," *Philosophy and Phenomenological Research* 69: 633–59.
- Sehon, S. R. (2000), "An Argument Against the Causal Theory of Action Explanation," *Philosophy and Phenomenological Research* 60: 67–85.
- Velleman, D. (1992), "What Happens when Someone Acts?" Reprinted in his (2000) *The Possibility of Practical Reason*, Oxford: Oxford University Press, pp. 123–43.
- Wilson, G. M. (1989), *The Intentionality of Human Behaviour*. Stanford: Stanford University Press.
- Yablo, S. (1997), 'Wide Causation'. *Philosophical Perspectives* 11: 251–81.