**Agent Causation and the Phenomenology of Agency**

Randolph Clarke

Several philosophers claim that the phenomenology of one’s own agency conflicts with standard causal theories of action, couched in terms of causation by mental events or states. Others say that the phenomenology is prima facie incompatible with such a theory, even if in the end a reconciliation can be worked out. Here it is argued that the type of action theory in question is consistent with what can plausibly be said to be presented to us in our experience of our agency. Several routes to a claim that there is nevertheless a prima facie incompatibility are examined, and all are found wanting. The phenomenology of agency, it is argued, is no threat to a standard causal theory of action.

There is, many philosophers and psychologists say, a distinctive phenomenology of one’s own agency, something that it is like to act. Some maintain that the phenomenology tells against a widely held theory of action, one on which actions count as such in virtue of being caused by—or causings by—mental events or states.[[1]](#endnote-1) Others hold that the phenomenology is prima facie incompatible with such a theory, though in the end a reconciliation can be worked out.

There are indeed difficulties faced by the kind of action theory in question.[[2]](#endnote-2) But I argue here that it is logically consistent with what can plausibly be said to be presented to us in our experience of our agency. I examine several routes to a claim that there is nevertheless a prima facie incompatibility, none of which, I argue, withstand scrutiny. As far as I can tell, a standard causal theory of action is not threatened by what it is like to act.

*1. The Phenomenology*

Writers who discuss the phenomenology of agency often distinguish between the experience of one’s *own* agency and that of (observing) the agency of *others*. Most of the attention is to the former, and it is on the character of this that I focus here.

Writers use a variety of expressions to indicate what they mean by the phenomenology in question, including ‘the experience of acting’ (Bayne 2008: 184), ‘agentive experience’ (Bayne 2008: 184), ‘the experience of first-person agency’ (Bayne and Levy 2009: 50), ‘the experience of oneself as being active’ (Nida-Rümelin 2007: 258), ‘a sense of agency, of oneself as the agent of action’ (Marcel 2003: 50), ‘one’s feeling of oneself as agentive’ (Marcel 2003: 51), ‘one’s awareness of one’s actions…*as actions* and *as one’s own*’ (Marcel 2003: 51), ‘what it is like to act in a certain way’ (Horgan and Tienson nd: 3), ‘the “something it is like” to experience oneself as behaving in a way that constitutes action’ (Horgan 2007: 185) and ‘the phenomenology of doing’ (Horgan et al. 2003: 323; Pacherie 2008: 195). I take it that despite the variety of expression, there is a common subject matter here.

It is commonly held that the phenomenology in question is *intentional*: the experience has an intentional object, it is about something. The intentionality is usually said to be a matter of *representation*: the experience has representational content, it presents the world as being a certain way. Although the content is generally thought to be at least in part non-conceptual, the experience is said to have satisfaction conditions, which are commonly taken to be *veridicality conditions*: when and only when the satisfaction conditions are met, the experience is veridical.[[3]](#endnote-3) The phenomenology of agency, on this view, can be illusory: it can represent things as being a certain way when in fact they aren’t that way.

How, then, is the representational content of the experience of one’s agency best expressed? Writers identify several aspects or components, only one of which is my primary concern here.[[4]](#endnote-4) As Timothy O’Connor puts it, ‘It does not seem to me (at least ordinarily) that I am caused to act by the reasons which favor doing so; it seems to be the case, rather, that *I* produce my decision *in view of* those reasons’ (1995: 196). Similarly, Terry Horgan states: ‘Agentive phenomenology presents one’s behavior to oneself as caused *by oneself*—and does not present it as being caused by *mental states* of oneself’ (2007: 184). Martine Nida-Rümelin holds that ‘to experience oneself as active in one’s doing can be described…by saying that we experience ourselves as *the cause* or *a cause* of what happens’ (2007: 262).[[5]](#endnote-5) Derk Pereboom (2015: 280) seconds the characterizations offered by Horgan and Nida-Rümelin.

Some of these writers take this aspect of the phenomenology to count against theories of action couched in terms of causation by mental events or states. (For brevity, I shall often say just *events* when referring to this disjunction of categories.) Of course, it is recognized that the phenomenology, and this aspect in particular, might be systematically illusory; but, the idea seems to be, we should have a defeasible preference for a theory on which this is not so. As Nida-Rümelin observes, ‘We should be reluctant to accept a theory that implies massive, fundamental and permanent error in the way we experience ourselves and others’ (2007: 261). In this spirit, O’Connor takes the phenomenology to support a view on which *free* actions, at least, are causings by agents and not by events or states.[[6]](#endnote-6) Nida-Rümelin and Pereboom take it to provide support for similar views of *action in general*. (They differ on an important detail of their views, with Nida-Rümelin taking actions to be *agent causings of events*, while Pereboom takes actions to be *agent caused events*.) Although eventually arguing that the content of agentive experience is compatible with a view of our behavior as caused by mental states, Horgan maintains that

agentive phenomenology provides prima facie reason to think that a prerequisite for genuine agency is that the behaviors we experience as actions are not caused by prior states—not by states of the external environment, not by internal states of the agent, and not by some combination of the two. (2007: 184)

Although a theory of action that appeals to causation by agents is, of course, a causal theory, here I shall use ‘causal theory of action’, as many theorists do, to indicate a view on which actions count as such because they are caused by—or causings by—mental events or states. (I do *not* take the kind of theory in question to be committed to a reductive analysis of action-sentences in terms of event or state causation; the distinction is explained in section 3.) The question raised, then, is whether the phenomenology of first-person agency tells against, or is at least prima facie incompatible with, an action theory of this kind.

*2. Things and Causes*

We commonly make and accept claims on which a certain object or some volume of stuff is said to cause something.[[7]](#endnote-7) The Sun, we say, causes the skin condition that we call sunburn. Some water, we might say, dissolved a bit of sugar. When that happened, we may accept, the water caused the sugar to dissolve.[[8]](#endnote-8)

About the latter case, some will say that, strictly speaking, it was some *event* involving the water—its coming into contact with the sugar, perhaps—or some *state* of the water—perhaps its solvency—that caused the effect. (Similarly, mutatis mutandis, for the first case.) Perhaps some will add a further claim, namely, that statements citing objects or volumes of stuff as causes are in every case false. We commonly make and accept them, but when we make them we speak falsely, and when we accept them we accept false claims.

However, there is an alternative that does not convict so much of what we think and say of falsehood. Theorists who find it advantageous to avoid convicting the phenomenology of agency of systematic illusion might, for a related reason, favor the alternative here.

Causal claims citing objects or quantities of stuff as causes might be said to be in each case semantically equivalent to claims citing events as causes. The purported equivalence might be advanced in furtherance of a reductive analysis of the former to the latter. Employing an ordinary, non-technical notion of substance on which things like the Sun and a volume of water count, consider the following proposal as purporting to reflect an analysis of this kind:

Substance *s* caused event *e2* just in case there was some event, *e1*, such that *e1* involved *s* and *e1* caused *e2*.[[9]](#endnote-9)

One kind of event involving a substance is a change undergone by that substance. Understood in this light, the proposal implies that the statement that a certain volume of water caused some sugar to dissolve is equivalent to a statement that there was some change undergone by the water that caused this effect. (The latter statement does not identify the change, but it might have been the water’s coming into contact with the sugar.) If the equivalence is correct, then when some change undergone by a volume of water causes some sugar to dissolve, the water causes the sugar to dissolve.

A reductive analysis will claim that the right-hand side of the schema has a *conceptual priority*: it provides an analysis of the substance-causal concept employed in the left-hand side. Substance causation is thus said to lack conceptual fundamentality. An *ontological* reduction will claim, additionally, that in any given case, the instance of the right-hand side tells us *what the substance causation consists in*. Event causation is then said to be *ontologically prior* to causation by substances. The latter is thus said to lack ontological fundamentality.

Semantic inquiry cannot provide reasons for making this further ontological claim. If there are grounds for the claim, presumably they come from consideration of the metaphysics of causation. A proponent of the ontological reduction might argue that such considerations yield the view that the fundamental truthmakers of true causal statements citing substances as causes are causal relations having events as their first relata. (A similar line might be followed for a view of substance causation as reducible, instead, to causation by states of substances.)

If something like the reductive analysis sketched here is correct, then we would go wrong in inferring from the fact that sunburn is caused by emissions of ultraviolet radiation from the Sun that it is not caused by the Sun. On the contrary, given the reduction, if the first of these things is true, then so is the second.

I do not claim that substance causation is in this way reducible. Rather, my point here is that the logical possibility of such a reduction shows that the content of an experience presenting a substance as causing something—provided that it does not further present that causation as irreducible—is logically consistent with a view on which, fundamentally, *all* causation is by events. Why think, then, that there is even a prima facie incompatibility between agentive phenomenology that presents our actions as caused by, or causings by, ourselves and a theory on which actions are caused by, or causings by, mental events (or states)?

*3. Agent Causation and Reduction*

An agent, we might say, is any being who acts. And we might then count any causation by such a being (if there is any) as an instance of agent causation. But ‘agent causation’ is commonly used in a more restricted sense, to refer not just to *any* causation by a being who is an agent, but specifically to *manifestations of an agent’s capacity to act*. Hornsby (2010), for example, takes causation by an agent to be an agent’s exercise of such a capacity; and Nida-Rümelin (2007) appeals to agent causation only with respect to ‘doings’ in which a subject of experience is active. If a being who is an agent might cause a disturbance by, say, falling against a table set with fragile crystal, we do not have in that case agent causation in the restricted sense. We have this only when an agent acts.

If we take it that when an agent acts, the agent causes something in the restricted sense just identified, then a reductive analysis of action in terms of event causation would give us a reductive analysis of such agent causation. Some proponents of a causal theory of action maintain that such an analytic reduction can indeed be provided, and, of course, some candidate analyses have been offered. If any such proposal succeeds, then a representation of agency as involving agent causation does not tell against a causal theory of action, provided that it does not further present agent causation as irreducible.

Roderick Chisholm, a prominent proponent of agent causation as an ontologically fundamental phenomenon, apparently would have agreed. He wrote:

The issues about ‘agent causation’…have been misplaced. The philosophical question is not—or at least it shouldn’t be—the question whether or not there is ‘agent causation.’ The philosophical question should be, rather, the question whether ‘agent causation’ is reducible to ‘event causation.’ (1978: 622)

As I argue in the next section, the range of views on which agent causation is affirmed but held not to conflict with a causal theory of action is broader than what is recognized in this passage by Chisholm. It includes views on which reduction is said not to be possible.

The phenomenology of first-person agency, it is held, presents one’s actions as caused by, or causings by, oneself, and it does not present them as caused by, or causings by, mental events. But of course, *not* presenting something as having a certain feature is not equivalent to presenting that thing as *lacking* that feature. The phenomenology might *fail* to present one’s actions as caused by, or causings by, events without presenting them as *not* caused by, or causings by, such things.[[10]](#endnote-10) If agent causation is reducible in the indicated way, simply failing to present one’s actions as caused by, or causings by, events is not presenting anything incompatible with a causal theory of action.

Note that I have not made the following claim: ‘the experience of the self as a substantival cause [in one’s actions] is parasitic on, and perhaps reducible to, the experience of event (or state) causation’ (Bayne 2008: 194).[[11]](#endnote-11) I have not suggested that *the experience* of agent causation might be reducible. Rather, I have raised the possibility that *agent causation* is reducible.

Horgan (forthcoming) considers a challenge to a causal theory of action that begins:

If the behaviors I experience as my actions are really state-caused by mental states of myself like occurrent thoughts and occurrent wishes (or by combinations of such states), then these behaviors are not really produced by me myself.

To the extent that reductive analysis is an epistemic possibility, this line of thought lacks plausibility.

*4. Non-Reductively Realized Agent Causation*

It has long been recognized that there can be causation of bodily motions by agent-involving mental events of the sort favored by causal theories of action without yielding an instance of action. The causal process from such items to bodily motion might be wayward or deviant, running via states of nervousness or the intentions of other agents.[[12]](#endnote-12) A reductive analysis of agent causation, in its restricted sense—as the manifestation of a capacity to act—would have to provide, a priori, a necessary and sufficient condition for action, without resort to an unanalyzed notion of agency or agent causation, that would rule out wayward causation of the problematic sort. It is a contested matter whether any such analysis is possible.

John Bishop, despite holding that the problem can be solved, nevertheless denies that a causal theory of action provides a *conceptual reduction* of agent causation.[[13]](#endnote-13) Still, he remarks, to refute the claim that agent causation is *ontologically* fundamental,

it will be enough to show that, even if to act does not *mean* to be caused to behave by one’s reasons, nevertheless, action can exist without the need for special agent-causal relations in nature *because actions are realized within complex sequences of events related only in event-causal ways to one another*. And CTA [the causal theory of action], of course, suggests just what sort of complex, causally related event-sequences do realize actions, namely, certain kinds of sequences in which the agent’s mental states cause his or her behavior. Thus, agent-causation might be conceptually essential but not ontologically essential for action—and the latter is all we need to defend if we are to establish CTA. (1989: 96)

Certain event-causal processes might suffice for agent causation, and thus for action, because action is fully realized in them, even if conceptual reduction is not possible.

If this suggestion is tenable, then the suggested view is available also to theorists who, unlike Bishop, do not think that a necessary and sufficient condition can be given, a priori, that rules out causal deviance. For they as well may hold that, although reductive analysis is not possible, in each instance of action there are event-causal processes that suffice for, because they fully realize, action.

Bishop did not explain what it would be for action to be *realized* in event-causal processes, despite not being reductively analyzable. Fleshing out the idea would provide a way of understanding how agent causation, in the restricted sense—as the manifestation of a capacity to act—might fail to be ontologically fundamental even if it is not reducible.

Realization is commonly held to be an asymmetric (and hence irreflexive) relation of ontological dependence and determination. It is thought to allow for multiple realizability: any instance of the realized phenomenon requires some instance of the realizing phenomenon to realize it, but different kinds of realizers can realize instances of the same kind of realized phenomenon. The latter multiplicity stands in the way of any easy non-disjunctive specification of a condition that is both necessary and sufficient for the realized phenomenon in terms of the realizers.

A case that action, and hence agent causation in the restricted sense, is realized in causation by events might be made by establishing that the relation of the former to the latter has one or more of several features that are characteristic of realization.[[14]](#endnote-14) First, action, when present, might be said to be present *in virtue of* the presence of an instance of a certain kind of event-causal process. Second, an instance of action might be said to *consist in* the particular process of causation by events that suffices for it. (Indeed, there might be said to be *an identity of token instances* here, with a token action—and hence a token instance of agent causation—said to be identical with a token event-causal process.)[[15]](#endnote-15) Third, a causal process’s being some specific kind of process of causation by events that suffices for action might be said to be *a way of being* an instance of agent causation; the relation might be said to be that of *determinate* to *determinable*.[[16]](#endnote-16) Multiple realizability, in this last case, will be a matter of there being other kinds of processes of causation by events each of which is a determinate of the determinable *agent causation*.

Again, if there is warrant for these claims, it will presumably come, at least in part, from consideration of the fundamental metaphysics of causation. Such consideration provides this warrant if, for example, it supports the view that all causal phenomena are realized in causation by events. But if we have this warrant, then although affirming agent causation, we can fairly say that any given instance of it is nothing over and above a certain instance of causation by events—it is no addition of being.

Again, we see that if the phenomenology of first-person agency presents our actions as caused by, or causings by, ourselves, it does not thereby present us with anything logically inconsistent with a causal theory of action. Nida-Rümelin remarks that ‘it is desirable for any philosophical theory to be compatible with the content of…experiences [of agency]’ (2007: 255). Unless it is impossible for agent causation to be reducible to or realized in causation by events, an experience of our actions as caused by, or causings by, ourselves is compatible with a causal theory of action.[[17]](#endnote-17)

*5. Agential Experience and Metaphysical Neutrality*

The idea of agent causation that is ontologically fundamental is one of significant metaphysical sophistication. It is the idea of causation of actions, or of events intrinsic to actions, by substances who are agents, and of such causation as not reductively analyzable in terms of causation by events, and, further, as not realized in causation by events. I conjecture that very few agents have this idea.[[18]](#endnote-18)

Might one’s experience of one’s own agency represent it as involving agent causation, *with the agent causation represented as fundamental*, even though one lacks the indicated concept? It *is* commonly thought that the representational content of our experience of our agency is at least in part non-conceptual. Still, it strikes me as incredible that non-conceptual representation, uninfluenced by possession of the concept in question, might draw the required distinction between agent causation that is not ontologically fundamental and agent causation that is.

There might, however, be human agents whose agentive phenomenology *does* present their own action as involving agent causation that is (represented as) fundamental. Some theorists maintain that experience can be subject to *cognitive penetration*. The representational content of an experience, it is held, can be altered by one’s grasp of certain concepts, one’s beliefs, one’s desires, one’s thoughts of various sorts. The visual experience of an expert forester, for example, might present a certain tree as a pine tree, while the visual experience of a similarly situated naïve viewer might lack that representational specificity.[[19]](#endnote-19) It is a matter of controversy whether such cognitive penetration is possible and, if so, exactly what it comes to.[[20]](#endnote-20) But assuming the possibility, it might be that some human agents who possess the necessary metaphysical expertise do indeed experience their actions as involving agent causation that is (represented as) fundamental, while most agents—even most adult human agents—do not.

I offer two comments. First, on this hypothesis, the conflict between the agential phenomenology of these few agents and a causal theory of action originates in *extra-experiential* representations, even if these representations subsequently influence the content of agential experience for these agents. It is not the phenomenology of first-person agency that is the source of the conflict. Second, and more importantly, on this hypothesis, the truth of a causal theory of action would not entail the *systematic* illusoriness of the experience of acting. For the illusion would arise for only a very few agents. Thus, a defeasible preference to avoid a theory of action that would commit us to the systematic illusoriness of the phenomenology of first-person agency would not count against a causal theory of action.

*6. Two Routes to a View of Agent Causation as Fundamental*

Suppose that our agential experience presents our actions as caused by, or as causings by, ourselves, without presenting them as involving agent causation that is (presented as) ontologically fundamental. Its representational content does not draw the metaphysical distinction; it is metaphysically neutral in this respect. But suppose, further, that action, and hence the agent causation involved in it, is not in fact reducible to or realizable in event causation. Then there is, after all, an incompatibility between the phenomenology and a causal theory of action.

Might we find an argument here that the experience of our agency tells against a causal theory of action? I shall consider two lines of argument that might be advanced in support of the claimed fundamentality. In both cases, the reducibility of action and its realizability in what are fundamentally event-causal processes are ruled out without appeal to phenomenology.

The first line of argument focuses specifically on agency. It argues, as many theorists do, that the problem of wayward causal processes precludes any reductive analysis of action. Further, it argues that no process of event causation is metaphysically sufficient for action, no such causal process can ‘amount to’ agent causation in the restricted sense.[[21]](#endnote-21) The latter is thus said to be fundamental, not reducible to or realizable in causation by events.

The portion of the argument concerned with wayward causation is a standard argument by counterexample: conditions offered as sufficient for action are shown not to suffice, or conditions said to be necessary are shown to be unnecessary. No appeal to agential phenomenology plays a role here. Further, since (as I’ve argued) agential phenomenology (at least in the case of most of us) does not draw the crucial metaphysical distinction, it presents no evidence bearing on the question of whether action, and hence the agent causation involved in it, can be realized in event-causal processes. The argument against realizability, then, gets no help from the phenomenology.

A second line of argument focuses not on agency but on the metaphysics of causation. It might argue, on the basis of such considerations, that fundamentally all causation is causation by substances.[[22]](#endnote-22) If that is correct, then at best a causal theory of action fails to show us what action is fundamentally. But again, the indictment of the causal theory would be complete without any appeal to the phenomenology of agency.

I am inclined to think that proponents of a view of agent causation as irreducible and unrealizable in causation by events would do better to pursue the second route. The idea that there is fundamental causation by a substance only in the case of agency strikes me as proposing an implausible metaphysical exceptionalism. If there is causation by substances that is not reducible to and not realizable in causation by events, then, I would expect, it exists throughout nature, constituting the activity of substances both animate and inanimate, macro and micro. It is plain that the phenomenology of our agency has no role to play in determining whether this is so.

*7. The Experience of Freedom*

There is no logical inconsistency between the content of our experience of our actions as caused by, or as causings by, ourselves and a causal theory of action. Why think there is even a prima facie incompatibility between the experience and the theory? It is not that action sentences are prima facie irreducible to statements concerning causation by events, or that action is prima facie unrealizable in such causation.[[23]](#endnote-23) One has to delve into these matters to find reasons favoring or disfavoring reducibility or realizability.

Horgan alleges a different source of the appearance of incompatibility, namely, *the appearance of freedom*. He gives the name *the hypothesis of the mental state-causation of behavior*—for short, the MSC hypothesis—to the claim that ‘normally the behaviors that one experiences as one’s actions are state-caused by certain psychological states of oneself’ (2007: 190). And here he calls the aspect of first-person agentive phenomenology that presents one’s behavior as caused by oneself *the self-as-source experience*. He writes:

Why is it prima facie plausible that the content of agentive phenomenology is incompatible with the MSC hypothesis, and with the other two…hypotheses [physical causal closure and determinism] as well? The virtually ubiquitous aspect of *freedom* in agentive phenomenology figures crucially, as an essential dimension of self-as-source experience. Experiencing one’s behavior as produced by oneself is fundamentally different from experiencing it as caused by internal *states* of oneself; and one key aspect of the experiential difference is that one does not experience the behavior as state-causally necessitated. Instead, one experiences the actually performed behavior as one among a range of alternative behaviors that are genuinely open to oneself—are real alternative possibilities that could be performed instead. But if the behavior that is thus experienced is nonetheless state-causally determined, then it is necessitated after all—even though it is not *experienced* as necessitated. Hence (one might well think), if the behavior is really state-caused, then it is not a piece of genuine action at all; the phenomenology of agency is illusory and non-veridical. The real source of one’s behavior is not really oneself, but instead is a state of oneself (or a combination of such states). (2007: 190-91)

The experience of one’s behavior as caused by oneself, Horgan says, is not distinct from an experience of oneself as acting freely, in the sense that it is up to oneself whether or not one does the thing one does (189). The latter is ‘an essential dimension’ of the former. And there is a prima facie incompatibility, Horgan alleges, between a causal theory of action and this experience of acting freely.

We may grant that there is a prima facie incompatibility between determinism and free will, even if we think that, in the end, the two can be shown to be compatible. However, a causal theory of action is not committed to determinism, nor to the less general thesis that when we act, our behavior is necessitated by the states that cause it (in the circumstances that obtain).[[24]](#endnote-24) Indeed, there are sophisticated libertarian accounts of free will that incorporate causal theories of action.[[25]](#endnote-25)

Horgan favors a reconciliation project that defends the compatibility of determinism (as well as the MSC hypothesis and physical causal closure) with free will. He makes a powerful case for the advantages of such a view. But it is open to a proponent of a causal theory of action to pursue the vindication of the phenomenology of agency via a different route, by developing a view requiring indeterministic causation by events at crucial points in processes leading to or constituting action.

There are serious difficulties confronting libertarian theories of free will that encompass causal theories of action. But, of course, every theory of free will, no matter what view of action it incorporates, faces serious difficulties.[[26]](#endnote-26)

Horgan says at one point that seeking to vindicate our belief in our agency by championing a libertarian account of free will ‘is in direct conflict with…defending the MSC hypothesis’ (2007: 192). The latter, recall, is the hypothesis that when we act our behavior is caused by certain of our mental states. But there is no logical conflict between libertarianism per se and this causal thesis, given that, as I observed, the thesis is compatible with indeterminism.

*8. The Experience of Event Causation*

While we experience our actions as freely performed, Horgan maintains, typically when we experience causation by events, ‘the cause *is* experienced as necessitating the effect (in the experienced circumstances)’ (2007: 189). An apparent incompatibility between the phenomenology of one’s own agency and a causal theory of action might be said to arise, then, given a conviction, based on one’s experience of event causation, that the latter is necessitation. For there is a prima facie incompatibility between free will and necessitation.

Typical causal processes are vulnerable to interference in vastly many ways. The collision of a rolling billiard ball with a stationary one causes the latter to roll away; but despite the collision, the effect would not have occurred if the stationary ball had been glued to the table, or if at the moment of collision an observer had reached out and picked it up, or if at that moment a meteor impact had pulverized it, or…. Causes and circumstances that jointly necessitate an effect must preclude all potential interferers, and it seems doubtful that we commonly experience conditions that in fact suffice to do so. If we nevertheless commonly experience causes and (experienced) circumstances as necessitating their effects (which I find doubtful), the experience would seem to be commonly non-veridical.[[27]](#endnote-27)

Note that, as well as experiencing our own agency, we also experience the efficacy of our agency. And this experience of efficacy need not be an experience of our agency as necessitating an outcome. If my free-throw attempt results in the basketball falling through the hoop, I might experience this result as caused by my toss but *not* experience the toss as necessitating the outcome (in the experienced circumstances).

Our experience of event causation need not be, and I suspect often is not, an experience representing causes as necessitating their effects, in the experienced circumstances. There is little if any support here for the allegation of prima facie incompatibility between our experience of our agency and a causal theory of action.

*9. An Incompatibility of Experiences*

It is sometimes said that not only *do we not* experience our actions as caused by, or as causings by, mental events, but that one *cannot* both experience something as an action one is performing and experience it as caused by, or as a causing by, events.[[28]](#endnote-28) Such a mutual exclusion of experiences might suggest an incompatibility of what is presented in them.

Horgan tackles this problem in light of his view that the experience of event causation presents it as necessitation, while the experience of one’s agency presents it as freely performed, and he suggests an evolutionary explanation for why the experiences are mutually exclusive.[[29]](#endnote-29) I shall briefly consider a suggestion that does not rely on the claim concerning necessitation.

Experiencing something as an action one is performing is experiencing it as its agent. In contrast, experiencing something as caused by, or as a causing by, events is typically experiencing it as a spectator of it. There is a difference here not just of *content* but of *perspective*: that of agent engaged in performing the action experienced, and that of spectator of some occurrence. This difference of perspective might explain a mutual exclusion of the experiences in question, even if their contents are compatible. For there seems to be some kind of incompatibility of the two perspectives.

In this light, we can see that the exclusion is limited. When I use a gardening tool (a lopper) to cut a limb whose thickness is at the limit of the tool’s cutting capacity, I may experience *the coming apart of the limb* both as caused by my exertions and as caused by me. *My action of cutting through the limb* might then be experienced as a causing of the outcome by me and as a causing of it by my exertions.

Where the incompatibility appears strongest is with respect to my exertions themselves. It is hard to imagine experiencing these as actions I am performing and, at the same time, experiencing them as caused by, or causings by, events. It is hard to imagine experiencing them at one and the same time from both of the two perspectives—that of agent engaged in making these exertions and as spectator observing them. What to make of this fact I am not sure; but it does not strike me as a grave threat to a causal theory of action.

We sometimes experience the behavior of others both as action and as caused by events. (Imagine observing orchestra members raise their instruments as the conductor raises the baton.) The limited exclusion, then, is evidently not due to an incompatibility between representation as action and representation as event caused.

*10. Conclusion*

For at least most agents, the representational content of the experience of first-person agency is not sufficiently sophisticated to distinguish between agent causation that is ontologically fundamental and agent causation that is non-fundamental, reducible to or realized in causation by events or states. Consequently, if when one acts one experiences one’s action as caused by, or as a causing by, oneself, the experience presents nothing logically inconsistent with a causal theory of action. Further, since such a theory is not committed to determinism, the experience of one’s actions as freely performed does not present a prima facie incompatibility, even given a prima facie incompatibility between determinism and free will. I considered several other routes to a claim of prima facie incompatibility, but all were found wanting. In sum, although proponents of a causal theory of action have their worries, it does not appear that the first-person experience of acting is one of them.

**Notes**

1. The more common version of such a theory follows Davidson (1980a, 1980c) in holding that when, for example, one raises one’s arm, the rising of the arm is one’s action of raising one’s arm, and it counts as an action because mental causes of the right sort cause it in the right way. An alternative (e.g., Bach 1980; Dretske 1988: ch. 1) takes actions to be *causings* rather than *events with certain kinds of causes*. For example, an act of raising one’s arm might be said to be a causing, in the right way, of the arm’s rising by mental events or states of certain sorts. Theorists who appeal to agent causation display a similar disagreement among themselves, with some (e.g. Pereboom [2014: ch. 3]) holding that agents cause actions and others (e.g. O’Connor [2000: 51]) holding that actions are agent-causings of events. [↑](#endnote-ref-1)
2. There is, for example, *the problem of mental causation*, the exclusion problem posed by the thesis of physical causal closure. There is *the problem of wayward causal processes* (see section 4). And there are various objections that are presented under the heading, *the disappearing agent* (see, e.g., Hornsby 2004a, 2004b). [↑](#endnote-ref-2)
3. Searle maintains that ‘the experience of acting just is the intention in action’ (1983: 91), which he takes to have a world-to-mind direction of fit. This would yield a view of the experience as *directive* rather than *descriptive* (as Bayne [2008: 187] puts it). Bayne (187) suggests a third alternative, on which the experience has both a mind-to-world and a world-to-mind direction of fit; borrowing an expression from Millikan (1996), he calls the experience on this view a *pushmi-pullyu* representation. He also mentions (though does not favor) a fourth option on which some or all of agentive experience involves raw phenomenal feels without intentional content (2008: 188). [↑](#endnote-ref-3)
4. In cases of overt bodily action, commonly one’s experience of agency includes perceptual and kinesthetic experience. In many cases of mental as well as overt bodily action, there is an experience of effort (Bayne and Levy 2009: 50). Other aspects that writers mention are an experience of purposiveness (Horgan 2007: 187), a sense of initiation (Pacherie 2007: 206), an experience of authorship (Bayne and Levy 2009: 50), and an experience of being able to do otherwise, or of its being up to oneself whether one does this or that (Bayne 2008: 195; Horgan et al. 2003: 331-32). This last is sometimes said to be an ‘essential dimension’ (Horgan 2007: 190) of the experience of one’s behavior as caused by oneself; I turn to this issue in section 7. [↑](#endnote-ref-4)
5. In a later paper, Nida-Rümelin (2018) backs away from the claim that in experiencing ourselves or others as active in behavior, we experience the behavior as *caused* by the agents in question, though she nevertheless maintains that we experience it as *brought about* by the agents. [↑](#endnote-ref-5)
6. O’Connor once (2001) held a view on which some actions (those in which free will is exercised) are agent-causings while others involve only event causation. He has since come to the view that all causes are substances (see Jacobs and O’Connor 2013). Note that, in arguing for the earlier view, the phenomenology of our agency was but one of several considerations (and not the primary one) that he offered in support of it.

   [↑](#endnote-ref-6)
7. Portions of this section and the next draw from Clarke (forthcoming). [↑](#endnote-ref-7)
8. This example presents what linguists call *causative alternation*. The verb ‘dissolve’ belongs to a family of verbs each of which has a transitive and an intransitive use. The alternation in this case is:

   a) The water dissolved the sugar.

   b) The sugar dissolved.

   It is commonly thought that ‘the transitive use of a verb V [in this family] can be paraphrased as roughly “cause to V-intransitive”’ (Levin 1993: 26-27). [↑](#endnote-ref-8)
9. Lowe (2008: 143) proposes this equivalence. It would, I think, require revision before yielding a credible analysis. For one thing, it should allow that when some substance causes something, some plurality of events involving that substance cause the effect.

   Lowe does not himself take the right-hand side of his proposed equivalence to provide a reductive analysis of the left-hand side. On the contrary, he takes the reduction to go in the other direction, with the fact presented in the left-hand side the more metaphysically fundamental one. More specifically, using ‘acting’ in a broad sense in which undergoing any change counts as acting, Lowe endorses the following reduction of event causation to substance causation:

   Event *e1* caused event *e2* just in case there was some substance, *s1*, and some substance, *s2*, and some manner of acting, *F*, and some manner of acting, *G*, such that *e1* consisted in *s1*’s *F*ing and *e2* consisted in *s2*’s *G*ing and *s1*, by *F*ing, caused *s2* to *G*. (2008: 145) [↑](#endnote-ref-9)
10. Horgan et al. (2003: 335) make a similar point. [↑](#endnote-ref-10)
11. Bayne suggests this contention, but he does not explicitly endorse it. In fact, Bayne and Levy express doubt that ‘in the usual run of things, the experience of first-person agency involves an experience of mental causation as such’ (2009: 52). [↑](#endnote-ref-11)
12. For examples, see Bishop (1989: 156-60) and Davidson (1980b: 79). [↑](#endnote-ref-12)
13. The theory that Bishop puts forward offers, a priori, a necessary and sufficient condition for action, which Bishop takes to involve agent causation. Why, then, does he deny that it provides a conceptual reduction? He appears to take the theory he advances to presume a ‘metaphysics of naturalism,’ and he apparently leaves open the possibility that there is action in worlds with other, non-natural ontologies. The proffered analysis might then be said to provide a reductive analysis of *natural action*, or action in worlds with the presumed ontology. On this point, see his 1989: 97. [↑](#endnote-ref-13)
14. In philosophy of mind, realization is most commonly taken to be a relation between properties, states, or events. The suggestion here is that it may also relate causings—instances of one thing’s causing another. Realization is often said to be not a single relation but a family of similar grounding relations. (See, for example, Baysan [2015].) The suggestion here is, then, that some member of this family might relate causation by events (or states) and causation by agents. [↑](#endnote-ref-14)
15. On these first three characteristics, see Kim’s discussion of realization in Kim (2005), esp. p. 579. [↑](#endnote-ref-15)
16. Yablo’s (1992) proposal that mental properties and events are related to physical properties and events as determinables to determinates offers a view of the mental as realized in the physical. [↑](#endnote-ref-16)
17. There is a rather different view that shows that the content of an experience of one’s agency as involving causation by oneself is logically consistent with a causal theory of action. The account of free will advanced by Clarke (1993) takes free actions to be caused by events *and* caused by agents. Unlike the view considered in the text here, that account takes agent causation not to be reducible to or realized in causation by events. The upshot of the discussion here is that logical inconsistency is avoided without such a commitment. [↑](#endnote-ref-17)
18. Indeed, it may be that even many who study the philosophy of action lack this idea. When agent causation is discussed, often the distinction that is drawn is between, on the one hand, causation by agents that is reducible to causation by events and, on the other hand, causation by agents that is not so reducible. Rarely is the possibility entertained of agent causation that is not so reducible but is nevertheless not ontologically fundamental. Perhaps philosophers of action commonly possess the latter idea, but their writing does not so often exhibit possession of it. [↑](#endnote-ref-18)
19. The example is from Siegel (2010). [↑](#endnote-ref-19)
20. See Raftopoulos and Zeimbekis (2015) for an overview of contemporary debates concerning the cognitive penetrability of perception. [↑](#endnote-ref-20)
21. Velleman sees a central task of a causal theory of action as that of explaining how causal relations between mental states and events of certain kinds and external behavior can ‘amount to a person’s causing something rather than merely to something’s happening in him’ (2000: 131). Hornsby responds: ‘They cannot’ (2000b: 186). [↑](#endnote-ref-21)
22. Advocates of this view include Jacobs and O’Connor (2013), Lowe (2008: 143-46), and Swinburne (1997, 2006). Alternatively, what might be argued for here is a causal pluralism on which entities of several categories, including substance, can be causes, with causation by none more fundamental than causation by the others. Hyman (2015: 40-42) and Steward (2012: 207-16) hold such pluralist views. [↑](#endnote-ref-22)
23. Every proposed reductive analysis of action faces difficulties concerning wayward causal processes. Does this fact itself generate a rational presumption in favor of the conclusion that action cannot be fully realized in processes of event causation? Compare the case of action with that of perception. Causal accounts of perception, like those of action, face difficulties concerning causal deviance. This fact does not itself create a rational presumption against the view that perception is fully realized in processes of causation by events. The case of action seems to me analogous in this respect. [↑](#endnote-ref-23)
24. It should be clear, as well, that a causal theory of action need make no commitment to what Horgan calls physical state-causal closure. [↑](#endnote-ref-24)
25. Some examples: Balaguer (2010: ch. 3); Ekstrom (2000: 99-129); Kane (1996); and Mele (2006: ch. 5). [↑](#endnote-ref-25)
26. It might be suggested that *the problem of luck* uniquely confronts an indeterministic theory of free will drawing on a causal theory of action, and thus presents us with a prima facie incompatibility of the experience of freedom with this kind of theory that does not arise with respect to competing theories. The problem, in brief, is that if on some occasion one’s action is caused but not determined by certain mental states of oneself, while there was a chance that instead one would act otherwise, with one’s doing otherwise caused by different mental states of oneself, then it seems to be just a matter of luck that one acts as one does on this occasion rather than acting otherwise. And if this is just a matter of luck, it seems, it can’t be that on this occasion one acts freely in doing what one does.

    However, the problem of luck confronts not only indeterministic theories that incorporate a causal theory of action, but also those that appeal to agent causation. (Mele [2006: ch. 3] and van Inwagen [2000] make this point.) Of course, it is sometimes argued by proponents of agent-causal theories that their views are uniquely capable of solving the problem of luck. But if that is so, argument is needed to establish that it is so. If the problem presents a prima facie incompatibility with the experience of freedom, the difficulty infects both kinds of view, even if only one, in the end, has the resources to fight off the infection.

    Pereboom (2014: 32) presses a ‘disappearing agent’ objection against event-causal libertarian theories, and it might be thought that this objection provides reason to think that the content of our experience of our actions (as, among other things, freely performed) is incompatible with a causal theory of action. Indeed, Pereboom (2015) advances such an incompatibility argument. I argue in Clarke (forthcoming) that the disappearing agent objection, as formulated by Pereboom, is ineffective. [↑](#endnote-ref-26)
27. Might such an experience, in those who have it, be due in part to cognitive penetration of a conviction that causation is necessitation (in the circumstances)? The question seems worth pursuing. [↑](#endnote-ref-27)
28. Horgan claims something slightly weaker than a strict exclusion here: ‘it is virtually impossible to simultaneously *experience* a single item of one’s own behavior both as actional and as state-caused’ (2007: 197). [↑](#endnote-ref-28)
29. The explanation that Horgan (forthcoming) offers notes the contribution to fitness that accrues from a capacity to distinguish between instances in which one is able to intervene and alter the course of events to one’s advantage and instances in which one is not able to so act. The idea, it seems, is that in the latter cases, but not in the former, one experiences certain events as necessitated by events that cause them (given the circumstances). The experience of oneself as able to act advantageously, it is suggested, thus excludes the experience of outcomes as event caused.

    But note that the latter group of cases includes some in which another agent poses a threat that one sees oneself as unable to resist but able to escape. In drawing the distinction in question, one need not see the conduct of the other agent as necessitated.

    **References**

    Bach, K. (1980). ‘Actions are not Events.’ *Mind* 89: 114-120.

    Balaguer, M. (2010). *Free Will as an Open Scientific Problem*. Cambridge, Mass.: MIT Press.

    Bayne, T. (2008). ‘The Phenomenology of Agency.’ *Philosophy Compass* 3: 182-202.

    Bayne, T., & Levy, N. (2009). ‘The Feeling of Doing: Deconstructing the Phenomenology of

    Agency.’ In N. Sebanz & W. Prinz (Eds.), *Disorders of Volition* (pp. 49-68).

    Cambridge, Mass.: MIT Press.

    Baysan, U. (2015). ‘Realization Relations in Metaphysics.’ *Minds and Machines* 25: 247-260.

    Bishop, J. (1989). *Natural Agency: An Essay on the Causal Theory of Action*. Cambridge:

    Cambridge University Press.

    Chisholm, R. M. (1978). ‘Comments and Replies.’ *Philosophia* 7: 597-636.

    Clarke, R. (1993). ‘Toward a Credible Agent-Causal Account of Free Will.’ *Noûs* 27: 191-203.

    Clarke, R. (Forthcoming). ‘Free Will, Agent Causation, and “Disappearing Agents”.’ *Noûs*.

    Davidson, D. (1980a). ‘Actions, Reasons, and Causes.’ In *Essays on Actions and Events* (pp. 3-

    19). Oxford: Clarendon Press.

    Davidson, D. (1980b). ‘Freedom to Act.’ In *Essays on Actions and Events* (pp. 63-81). Oxford: Clarendon Press.

    Davidson, D. (1980c). ‘The Logical Form of Action Sentences.’ In *Essays on Actions and Events* (pp. 105-122). Oxford: Clarendon Press.

    Dretske, F. (1988). Explaining Behavior: *Reasons in a World of Causes*. Cambridge, Mass.: MIT

    Press.

    Ekstrom, L. W. (2000). *Free Will: A Philosophical Study*. Boulder, Co.: Westview.

    Horgan, T. (2007). ‘Mental Causation and the Agent-Exclusion Problem.’ *Erkenntnis* 67: 183-

    200.

    Horgan, T. (forthcoming). ‘Causal Compatibilism about Agentive Phenomenology.’ In T.

    Horgan, M. Sabates, & D. Sosa (Eds.), *Supervenience in Mind*. Cambridge, Mass.: MIT

    Press.

    Horgan, T., & Tienson, J. (nd). ‘The Phenomenology of Embodied Agency.’ Available at

    www.u.arizona.edu/~thorgan/papers/mind/PhenomenologyEmbodiedAgency.doc.

    Horgan, T., Tienson, J., & Graham, G. (2003). ‘The Phenomenology of First-Person Agency.’ In

    S. Walter & H.-D. Heckmann (Eds.), *Physicalism and Mental Causation: The*

    *Metaphysics of Mind and Action* (pp. 323-340). Exeter, UK: Imprint Academic.

    Hornsby, J. (2004a). ‘Agency and Actions.’ In J. Hyman & H. Steward (Eds.), *Agency and Action* (pp. 1-23). Cambridge: Cambridge University Press.

    Hornsby, J. (2004b). ‘Agency and Alienation.’ In M. De Caro & D. Macarthur (Eds.), *Naturalism in Question* (pp. 173-187). Cambridge, Mass.: Harvard University Press.

    Hornsby, J. (2010). ‘The Standard Story of Action: An Exchange (2).’ In J. H. Aguilar & A. A.

    Buckareff (Eds.), *Causing Human Actions: New Perspectives on the Causal Theory of*

    *Action* (pp. 57-68). Cambridge, Mass.: MIT Press.

    Hyman, J. (2015). *Action, Knowledge, and Will*. Oxford: Oxford University Press.

    Jacobs, J. D., & O’Connor, T. (2013). ‘Agent Causation in a Neo-Aristotelian Metaphysics.’ In

    S. C. Gibb, E. J. Lowe, & R. D. Ingthorsson (Eds.), *Mental Causation and Ontology* (pp. 173-192). Oxford: Oxford University Press.

    Kane, R. (1996). *The Significance of Free Will*. New York: Oxford University Press.

    Kim, J. (2005). ‘Supervenience, Emergence, Realization, Reduction.’ In M. J. Loux & D. W.

    Zimmerman (Eds.), *The Oxford Handbook of Metaphysics* (pp. 556-584). Oxford: Oxford

    University Press.

    Levin, B. (1993). *English Verb Classes and Alternations: A Preliminary Investigation*. Chicago: University of Chicago Press.

    Lowe, E. J. (2008). *Personal Agency: The Metaphysics of Mind and Action*. Oxford: Oxford

    University Press.

    Marcel, A. (2003). ‘The Sense of Agency: Awareness and Ownership of Action.’ In J. Roessler

    & N. Eilan (Eds.), *Agency and Self-Awareness: Issues in Philosophy and Psychology* (pp.

    48-93). Oxford: Oxford University Press.

    Mele, A. R. (2006). *Free Will and Luck*. New York: Oxford University Press.

    Millikan, R. G. (1995). ‘Pushmi-Pullyu Representations.’ *Philosophical Perspectives* 9:

    185-200.

    Nida-Rümelin, M. (2007). ‘Doings and Subject Causation.’ *Erkenntnis* 67: 255-272.

    Nida-Rümelin, M. (2018). ‘Freedom and the Phenomenology of Agency.’ *Erkenntnis* 83: 61-87.

    O’Connor, T. (2000). *Persons and Causes: The Metaphysics of Free Will*. New York: Oxford University Press.

    Pacherie, E. (2008). ‘The Phenomenology of Action: A Conceptual Framework.’ *Cognition* 107: 179-217.

    Pereboom, D. (2014). *Free Will, Agency, and Meaning in Life*. Oxford: Oxford University

    Press.

    Pereboom, D. (2015). ‘The Phenomenology of Agency and Deterministic Agent Causation.’ In

    H. Pedersen & M. Altman (Eds.), *Horizons of Authenticity in Phenomenology, Existentialism, and Moral Psychology: Essays in Honor of Charles Guignon* (pp. 277-

    294). Dordrecht: Springer.

    Raftopoulos, A., & Zeimbekis, J. (2015). ‘The Cognitive Penetrability of Perception.’ In J. Zeimbekis & A. Raftopoulos (Eds.), *The Cognitive Penetrability of Perception: New Philosophical Perspectives* (pp. 1-56). Oxford: Oxford University Press.

    Searle, J. R. (1983). *Intentionality: An essay in the Philosophy of Mind*. Cambridge: Cambridge University Press.

    Siegel, S. (2010). *The Contents of Visual Experience*. Oxford: Oxford University Press.

    Steward, H. (2012). *A Metaphysics for Freedom*. Oxford: Oxford University Press.

    Swinburne, R. (1997). ‘The Irreducibility of Causation.’ *Dialectica* 51: 79-92.

    Swinburne, R. (2006). ‘Relations between Universals, or Divine Laws?’ *Australasian Journal of*

    *Philosophy* 84: 179-189.

    Van Inwagen, P. (2000). ‘Free Will Remains a Mystery.’ *Philosophical Perspectives* 14: 1-19.

    Velleman, J. D. (2000). ‘What Happens When Someone Acts?’ In *The Possibility of Practical Reason* (pp. 123-143). Oxford: Clarendon Press.

    Yablo, S. (1992). ‘Mental Causation.’ *Philosophical Review* 101: 245-280. [↑](#endnote-ref-29)