

Responsibility, Agency, and the Disappearing Agent Objection

Derk Pereboom, Cornell University

Forthcoming in Le Libre-Arbitre, approches contemporaines, Jean-Baptiste Guillon (ed.),
Paris, Collège de France
Penultimate draft

I've argued that the objection that reveals the deepest problem for event-causal libertarianism is what I call *the disappearing agent argument* (Pereboom 2004, 2014, 2017). This objection has recently been the target of several illuminating criticisms, and in what follows I respond to them, thereby clarifying the nature of the argument. Other arguments against event-causal libertarianism include the present luck objection (e.g., Haji 2004, Mele 2006), and the objection that the event-causal view does not yield enhanced control relative to that available given causal determination (e.g., O'Connor 2000, Pereboom 2001, Clarke 2003). I will explain why I believe that the disappearing agent argument is the strongest of the three, and is best at focusing our attention on the core concern for event-causal libertarianism.

This objection is related to a similar disappearing agent argument that targets not just event-causal theories of moral responsibility, but any event-causal theories of action, whether indeterministic or deterministic (e.g., Velleman 1992, Hornsby 2004, Steward 2012). I've said that I am cautiously sympathetic to this more general argument (Pereboom 2014, 2015), but that I don't need to endorse it to affirm the disappearing agent argument against event-causal libertarianism (Pereboom 2014, p. 32, note 1). But Al Mele (2015) and Laura Ekstrom (2016) ask whether endorsing the less general argument commits me to the more general one. I will explain my position on this issue, a key to which is that I believe the more general argument to be significantly weaker.

1. Three kinds of libertarianism

Three kinds of libertarianism have been proposed, distinguished by the causal story they tell: event-causal, agent-causal, and non-causal libertarianism. In event-causal libertarianism, actions, conceived as agent-involving events – as agents acting at times – are caused solely by prior events, such as an agent's having a desire or a belief at a time, and some type of indeterminacy in the production of actions by appropriate events is held to be necessary for the action to be free in the sense required for moral responsibility (e.g., Kane 1996; Ekstrom 2000; Balaguer 2010; Franklin 2011). In my view, the notion of moral responsibility at stake in the free will debate is the one that involves basic desert.¹ I am

¹ For an agent to be morally responsible for an action in the basic desert sense is for the action to be hers in such a way that she would deserve to be blamed if she understood that it was morally wrong, and she would deserve to be praised if she understood that it was morally exemplary. The desert at issue here is basic in the sense that the agent, to be morally responsible, would deserve to be blamed or praised just because she has performed the

counting basic actions such as decisions, and more generally, the formation of intentions, as actions. In theories that I will classify as within the event-causal libertarian camp, actions are indeterministically caused by states or property instances. This type of position has an early representative in the Epicurean view according to which a free decision is an indeterministically caused swerve in the otherwise downward path of an atom (Lucretius 50 BCE).

In agent-causal libertarianism, free will of the sort required for moral responsibility is accounted for by the existence of agents who fundamentally as substances have the power to cause actions without being causally determined to do so (e.g., O'Connor 2000, 2008; Clarke 2003). The causation involved in free action is not reducible to causation among events, and what ensures this is that the agent, fundamentally as a substance, has the power to cause an action. For an action to be free it's crucial that the agent who causes the action is not causally determined to cause it.

The third kind of libertarianism is non-causal (e.g., Bergson 1889/1910; Ginet 1997, 2007; McCann 1998; Goetz 2008). Early non-causal views arose in the context of a Kantian framework, according to which 'causation' is a concept we impose on spatial and temporal experience, and its applicability to some phenomenon requires that it be law-governed. Henri Bergson, an early non-causalist, contends that although actions occur in time, the time of conscious agency does not feature the magnitudes that mark the applicability of causal laws. Any attempt to theorize about conscious agency will involve invoking causal concepts, which are derived from physical theory, but do not in fact apply and are at best metaphorical. Conscious agency is *sui generis*, it is not subject to theorizing and is not causal in nature. This independence of causality, on Bergson's account, allows actions to be free. Contemporary non-causalists don't invoke this Kantian framework, at least not explicitly. In Carl Ginet's view, for a basic action to be free it must have an agent as a subject, it cannot be caused, and it must have an actish phenomenological feel. In place of the actish phenomenological feel, Hugh McCann specifies that the action be intentional, and intrinsically and fundamentally so, so that an action's being intentional is not in turn to be analyzed causally. On Stewart Goetz's conception, the action must satisfy a teleological requirement, and since McCann's intrinsic intentionality is teleological as well, their positions are similar in one key respect.

A significant difference between the non-causal view on the one hand and the event- and agent-causal views on the other is that for the non-causalist *control in action* is not, or at least not necessarily, a matter of the action's being caused by the agent or by agent-involving events. For the other views, control is necessarily a causal matter. This difference is important to the discussion that follows.

If the disappearing agent argument undermines event-causal libertarianism, libertarians may turn the agent-causal or to the non-causal version instead. In my view, neither of these alternatives falls to the disappearing agent argument, and each yields a response to this objection.

action, given sensitivity to its moral status; and not, for example, in virtue of consequentialist or contractualist considerations. (Pereboom 2014, 2001; cf. Feinberg 1970)

2. Disappearing agents

Here is the disappearing agent objection to event-causal libertarianism (the statement I currently prefer):

Suppose that a decision is made in a deliberative context in which the agent's moral motivations favor deciding to A, her prudential motivations favor her deciding to not-A, and the strengths of these motivations are in equipoise. A and not-A are the options she is considering. The potentially causally relevant events thus render the occurrence of each of these decisions equiprobable. But then the potentially causally relevant events do not settle which decision occurs, that is, whether the decision to A or the decision to not-A occurs. Since, given event-causal libertarianism, only events are causally relevant, *nothing* settles which decision occurs. Thus it can't be the agent or anything about the agent that settles which decision occurs, and she therefore lacks the control required for moral responsibility for it. (Pereboom 2004, 2014, 2017; cf. O'Connor 2008; Griffith 2010)

I call this objection the 'disappearing agent argument' because the agent (metaphorically) vanishes at the point at which we expect that she can settle whether the decision occurs. (Randolph Clarke (2017) suggests that a better moniker might be the 'fading agent objection,' since I'm open to the claim that agency is preserved in such situations.) The concern is that because event-causal libertarian agents do not have the power to settle which decision occurs, they cannot have the role in action that secures the control that moral responsibility demands. I use the equipoise situation only as a paradigm case. Event causal libertarians generally accept freedom of choice in non-equipose situations. The concern about settling arises in such contexts as well.

Al Mele (2015) asks for clarification on what it is for an agent to settle whether a decision occurs. I reply as follows (Pereboom 2017): First, to settle whether to A or to not-A, or equivalently, which of A and not-A occurs, is to *determine*, not necessarily causally, which of these options occurs.² As a matter of the psychology of action, in the typical case what an agent aims to settle is not whether a decision occurs; the focus is instead on a bodily motion, such as my raising my arm (e.g., Steward 2012, 32-36; Runyan 2016, 1630), or a mental analogue, such as my imagining my childhood home. The agent then determines that she *decides to A* by determining *to A*; I determine that I decide to raise my arm by determining that I raise my arm. Mele writes: "in deciding to A, one settles upon A-ing (or upon trying to A), and one enters a state—a decision state—of being settled upon A-ing (or upon trying to A)" (Mele 1992: 158–59). As he points out, in some possible cases one's focus in deliberating is on whether to decide to A by contrast with whether to A. If the evangelist says: "The time to decide has come – will you decide to commit your life to Jesus Christ?" and you believe that your eternal fate depends on your decision, specifically, you may focus on whether to so decide. Mele is correct to think that this would be an unusual type of case. We don't typically focus on the decision, but on the action at which the

² Laura Ekstrom (2016) appropriately warns against a characterization of settling that requires it to be causally determining.

decision aims; we then settle whether the decision occurs by virtue of settling on the action.

There is more to the notion of settling than determining whether an action or decision to act occurs. I agree with Helen Steward (2012) that settling in addition involves *difference making* – the agent making the difference as to which option for action occurs. Putting the two together, I propose the following biconditional characterization:

An agent settles which option for action occurs just in case she determines, not necessarily causally, which action occurs, and she makes the difference as to which action occurs.

A leading issue at stake in the free will debate is whether we have the control in action required for moral responsibility in the basic desert sense. For an agent to settle which option for action occurs is for her to exercise a crucial sort of control in action, the sort on which our being morally responsible in this sense would have to depend. For the event-causal libertarian, control is exclusively causal matter. For such an event-causalist, then, settling will be a causal matter, and plausibly both the determination and difference-making aspects of settling will be causal as well.

Steward (2012) and Clarke (in correspondence, cf. 2017) suggest that the settling relation is a relation between a concrete particular such as an event or an agent, and a proposition. On this conception, the settling relation cannot be causal, since the relevant relation to propositions, paradigmatically *making-true*, is essentially non-causal. But this isn't my notion of settling; it's not the one that I invoke in my disappearing agent argument. Clarke (in correspondence) agrees that on his essentially non-causal notion, whether settling relations obtain may depend on some causal relation. My notion of settling might then pick out the causal relation that fixes whether Clarke's essentially non-causal relation of settling obtains.

One might ask: is such difference-making compatible with the agent's being causally determined to act by factors beyond her control? Steward (2012) thinks not. I disagree (Pereboom 2015). In a deterministic context it's still often true that if the agent's reasons had been different, she would have acted differently. Thus it's open to the determinist to invoke a reasons-responsive notion of difference making. One might in addition ask whether difference making is possible in Frankfurt (1969) examples. In such scenarios, an agent is intuitively morally responsible for A despite not being able to do otherwise. He cannot do otherwise due to a potential intervention, poised to cause her to do A if it's indicated that she might do otherwise. I'm a Frankfurt-defender, and I'm committed to the claim that such cases feature settling (Pereboom 2001, 2014). However, in Frankfurt examples it's false that if the agent's reasons had been different, she would have acted differently. But Carolina Sartorio (2013) persuasively argues that causes of action in Frankfurt cases still make a difference to their effects in that the effects wouldn't have been caused by the absence of their causes. In the non-actual alternative scenario this absence must be supplemented by the intervention for the action to occur. (If one is averse to absence causation, one might invoke causal explanation in terms of absences instead (Sartorio 2016).)

Following this lead, we can propose the following difference-making-capturing sufficient condition for settling that applies generally, including deterministic contexts and in Frankfurt examples:

(S-EC) An agent settles whether an action occurs if it is caused by certain reasons of hers, where the absence of those reasons would not have caused that action.

Adapted to an agent causal context, we get the following sufficient condition:

(S-AC) An agent settles whether an action occurs if she agent-causes it for certain reasons, where the absence of her agent-causing the action for those reasons would not have caused that action.³

To settle which option for action occurs is an exercise of control in action, and its characteristics are determination and difference making. Mele (2015) correctly specifies that this is an exercise of direct control in acting, by contrast, with, say, gathering evidence and weighing reasons concerning how to act.

3. Two deflationary responses

Mark Balaguer (2014) and Randolph Clarke (2017) propose what might be called *deflationary* responses to the disappearing agent objection. In developing his event-causal libertarianism, Balaguer specifies that the sort of freedom at issue requires authorship and control: “in order for a decision to be L-free, it has to be *authored* and *controlled* by the agent in question: that is, it has to be her decision, and she has to control which option is chosen.” (2010, 83) The contexts he highlights in his discussion feature a *torn decision*, one in which two options are in equipoise with respect to motivating reasons. To illustrate his account, and to facilitate objections and his answers to them, Balaguer sets out his example of Ralph, who makes a torn decision about whether to stay in town or move to the big city:

Ralph is deciding whether to stay in Mayberry or move to New York. Favoring the move to New York are his desire to play for the Giants, and his desire to star on Broadway. Favoring staying in Mayberry are his desire to marry Robbi Anna, and his desire to manage the local *Der Weinerschnitzel*. Suppose Ralph makes the torn decision to move to New York – he just decides to move to New York. (2010, 72)

Balaguer subsequently sets out a version of Peter van Inwagen’s (1983, 141) rollback objection to libertarian views, to which he provides a reply that he also invokes in response to the disappearing agent argument (Balaguer 2014):

Suppose that some agent S is torn between two options, A and B, and eventually chooses A in a torn-decision sort of way. And now suppose that God “rolls back” the universe and “replays” the decision... If the decision is undetermined in the manner of TDW-indeterminism, then if God “played” the decision 100 times, we should expect that S would choose A and B about 50 times each.... It seems to be a matter of

³ This discussion of difference making corrects for problems for the earlier account I provide in (Pereboom 2015). Thanks to Richard Holton and Carolina Sartorio for valuable discussion.

chance or luck what she chose, and to the extent that this is right, it seems that S didn't author or control the decision. At the very least, it seems that the element of chance or luck here diminishes S's authorship and control. (2010, 92)

Setting the rollback objection in the context of this example, Balaguer's deflationary response is that "in each of the different plays of the decision, it is *Ralph* who does the choosing" (2009, 93). More generally, in his view:

...the most we could hope for, vis-à-vis authorship and control is that it be *Ralph* who does the just-picking... Ralph chooses – consciously, intentionally, and purposefully – without being causally influenced by anything external to his conscious reasons and thought. Thus it seems that in this case, we do get the result that it is Ralph who does the just-choosing. And so it also seems that in this scenario, we procure as much authorship and control for Ralph as we can, given that he is making a torn decision. (2010, 97)

Later, in his response (Balaguer 2014) to my commentary (Pereboom 2013) on his (2010) book, Balaguer indicates that he endorses the same response to the disappearing agent argument.

Clarke (2017) also develops a deflationary reply to the disappearing agent argument. He first introduces a condition on settling that he believes the disappearing agent argument relies upon:

(SR) If an agent S freely decides at time t to A, then S settles at t whether that decision is made then.

Clarke joins Mele (2015) in objecting to my intimation that on the event-causal libertarian view, events antecedent to the decision would need to settle whether it *will* occur. I said:

On an event-causal libertarian picture, the relevant causal conditions antecedent to the decision, i.e., the occurrence of certain agent-involving events, do not settle whether the decision will occur (Pereboom 2014, 32).

Clarke's SR specifies, by contrast, that settling happens exactly when the decision is made, a view for which Christopher Franklin (2014, 425) also argues. Clarke justifies this specification as follows:

Standard libertarian accounts—including standard event-causal libertarian views—require that directly free decisions aren't determined by anything that precedes them. Proponents of such views, if they accept a settling requirement, will take it to require that when an agent exercises free will in making a decision, the settling takes place when the decision in question is made, not prior to that time.

These considerations motivate Clarke's deflationary reply:

The Rejoinder: the making of the decision by S at t to A settles whether the decision is made then.

He adds:

After all, that matter is not settled by anything prior to t (for the decision is not determined by anything prior to t); and nothing more than the making of the decision at t is needed to settle the matter then. Further, since it is S who makes the decision, S, in making that decision, settles at t whether that decision is made then. For given that nothing prior to t settles whether that decision is made then, S need not do anything more than decide at t to A in order to settle at t whether that

decision is made then. An event-causal libertarian theory, then, has the resources to satisfy the settling requirement SR.

Let me begin by responding to Clarke's deflationary reply. On the Rejoinder, the making of the decision by S at t to A settles whether the decision is made then:

To clarify, by 'the making of the decision', what is meant is simply the occurrence of the mental action of deciding. If such a mental action occurs, then a making of a decision takes place; the latter is nothing more and nothing less than—it *is*—the former. (Clarke 2017)

The proposal is that there is an identity between *the making of the decision* and *the occurrence of the mental action of deciding*, and that the occurrence of the mental action of deciding settles which decision occurs. Clarke's thought is that the occurrence of the mental action of deciding settles which decision occurs in the sense of making a certain proposition about the decision true (2017, and from correspondence). Now I agree that such a use of 'settling' is fine linguistically. But it's not what I mean by 'settling.' Settling, as I use the term in the context of the disappearing agent argument, is an exercise of control, and the event-causal libertarian is committed to this exercise of control being a causal matter. The relation between the occurrence of the mental action of deciding and whether the decision occurs is not a causal relation, and thus the Rejoinder is at odds with the principle that the control involved in settling is a causal matter.

How should we interpret "the making of the decision by S at t to A," given the event-causal view, and the supposition that this phrase picks out an exercise of control that is an instance of settling in the sense of 'settling' that I have in mind? The making of the decision by S would consist in the causing of the event *S's deciding to A at t* by S-involving events, say (desire) E1 and (belief) E2. In Balaguer's example, Ralph's just-deciding is caused by, say, his desiring to star on Broadway and his believing that moving to New York will facilitate this aim. But given indeterminism and equipoise, there will be other events, E3 and E4, set to cause the alternative action-event, *S's deciding to B at t*, in Ralph's case, deciding to stay in Mayberry, with the same antecedent probability. However, at this point there is nothing left to settle whether *S's deciding to A at t* by contrast with *S's deciding to B at t* occurs. Only E1-E2 and E3-E4 are candidates for this role, but they don't settle which decision-event occurs. This is the disappearing agent argument, and so far we've encountered no convincing event-causal libertarian response.

But Clarke (2017) develops an interesting idea that stands to advance the discussion. He notes that on one variety of the agent-causal view, decisions are agent-causings, which in turn might be analyzed as activations of an agent-causal power (DeRose 1993; O'Connor 2000, 51). On this picture, which decision occurs is not settled by the agent causing it, but instead by virtue of *its being an agent-causing*. Clarke then proposes that we adapt this suggestion to the event-causal view. In that case, the decision is an event-causing, and we can plausibly claim that its occurrence is settled by virtue of its being identical to the event-causing. This proposal appears to combine the specification that settling is a causal matter with the thought that which decision occurs is settled by virtue of an event or state with which it's identical.

However, unless more is said (and more may be said -- see below), this conception of the agent-causal position doesn't really preserve the specification that control is a causal matter. In this conception, the agent does not cause her causings, and thus the relation between the agent and the decision is in fact non-causal. On a natural interpretation of the specification that control is a causal matter the control that agent causation supplies is control *over what an agent causes*. A view on which agents don't cause decisions, but instead exercise control over decisions by virtue of their being agent-causings, is not in accord with this natural interpretation. I would place this view, and its event-causal analogue, in the non-causal camp.

But there is a related interpretation of the agent-causal position that genuinely preserves the specification that agents exercise control over what they cause, in particular in the case of decisions. When an agent acts freely, what she typically does most fundamentally is to cause an action, by contrast with causing a decision to perform that action. (This accords with Mele's suggestion about settling cited earlier. (Mele 1992, 158–59)) Further, it is a logical consequence of S agent-causing action A at t1 that the event *S's causing A at t1* occurs. It is a logical consequence of the fact that I agent-cause my raising my arm at t1 that the event *my agent-causing my raising my arm at t1* occurs. In addition, we can plausibly specify that here my agent-causing my raising my arm at t1 is identical to *my deciding to raise my arm at t1*. Now, to adapt a suggestion of Roderick Chisholm's (1971), we can propose that in agent-causing my raising my arm at t1 I also *bring about* the event *my agent-causing my raising my arm at t1* -- that is, I also bring about *my deciding to raise my arm at t1*. Since bringing about just is causing, I thus also cause my deciding. In this way, the agent-causalist can preserve the view that control over decisions is a kind of control over what an agent causes.

The event-causal analogy to this picture is as follows. S can exercise control over action A by virtue of S-involving events causing A. *S's deciding to A at t1* would in such cases be identical to the complex event *S-involving events causing A at t1*. In causing A at t1, the S-involving events bring about, and thus cause, the event *S-involving events causing A at t1*, which is *S's deciding to A at t1*. However, given motivational equipoise, the S-involving events don't settle whether A occurs, and thus don't settle whether the decision to A occurs. In this analogy, the causes of the decision are inherited from the action, and since the causes of the action don't settle whether the action occurs, they don't settle whether the decision occurs either. This view thus falls to the disappearing agent argument. Perhaps the general lesson is that the disappearing agent argument successfully targets any event-causal view on which control in deciding is control over what an agent causes.

Alternatively, one might just specify that "the making of the decision by S at t to A" and "Ralph's just deciding to move to New York" are to be interpreted non-causally. In accord with this suggestion, Balaguer (2014) proposes a response to the disappearing agent argument that features a combination of event-causal and non-causal relations. Resisting the agent-causal resolution, he says: "in this [equipoise] scenario, the event that settles which option is chosen is the conscious decision – i.e., the event with a me-consciously-choosing-now-phenomenology" (2014, 83). On one non-causal option, "the making of the decision by S at t to A" and "S's just deciding to A" amount to a being-the-subject-of relation between S and A together with, say, intrinsic intentionality. This option

might be considered on its own merits, but it is at odds with the view that control exercised in settling is a causal matter, a commitment of event-causal libertarianism, at least as it's usually set out.

My main objection to non-causalism is that its proponents, such as Ginet (1997, 2007) agree that in the case of a basic action the agent *makes the action happen* and *makes the difference as to when it will happen*, but making-happen and this sort of difference-making would seem to be causal relations (Pereboom 2014, 39-47). My concern is that despite the claim of non-causalism, the settling relation would seem to be causal after all. A genuinely non-causal view, on the other hand, faces the problem of accounting for the control in action required for responsibility in a non-causal way.

Clarke's and Balaguer's deflationary views, if interpreted in a resolutely event-causal way, are subject to the disappearing agent argument after all, given my notion of 'settling.' They might also be interpreted non-causally, but then they fall afoul of a key commitment of event-causal libertarianism, that control in action is a causal matter, when this is interpreted as the claim that in action agents exercise control over what they cause.

4. Agent versus Event Causation

Mele (2015) argues, plausibly, that on event-causal libertarian views it remains correct to say that *agents*, and not events, decide and make decisions. "[Sam] has and exercises the ability or power to sink free throws, and he sinks many of them. His intentions, beliefs, skills, and the like do not sink free throws – alone or in combination with one another. And that is no surprise, because they are not able to sink free throws." Similarly, in her response to the disappearing agent argument Laura Ekstrom (2016) writes: "it is something about the agent that settles which decision occurs: the agent exercises her will in making a decision for reasons. In the case of the writer, for instance, in making the decision to spend the afternoon working on her paper, she settles what to do with her afternoon." Both Mele and Ekstrom are resolute in their adherence to the event-causal libertarian protocol; neither wishes to shift to a non-causal or an agent-causal alternative. But Ekstrom does employ ordinary agent talk, by contrast with language that is explicitly event-causal, in response to the disappearing agent argument, and Mele seems to be suggesting that it helps.

To be sure, retaining ordinary agent talk is legitimate for the event-causal libertarian. The view is not committed to linguistic reform. But it is committed to the metaphysical thesis that all causation is event-causation. What is Mele then contending when he says that Sam's "intentions, beliefs, skills, and the like do not sink free throws"? It's not open to an event-causal theorist to hold that in the sinking of a free throw *Sam* is a causal influence distinct from Sam-involving events. It is indeed strange to say that a collection of events sinks a free throw, and we're not in the habit of speaking this way. But

what's at issue here isn't how we're used to speaking, but the metaphysical commitments of an event-causal theory of action (Pereboom 2017).⁴

An event causalist might endorse the semantic position according to which 'Sam sunk the free throw' can be true, but the causal relations that ground its truth are solely event-causal relations (Bishop 1989). I maintain that the disappearing agent argument succeeds against an event-causal libertarianism that endorses this semantic position. For given equipoise, once the event-causal relations that ground the truth of the ordinary agent talk are specified, the settling problem becomes evident. I also contend that the argument also successfully challenges a libertarian view according to which the agent is a substance cause of action, while all substance-causal relations reduce to event-causal relations. Given equipoise, once the reduction base is specified, the settling problem again becomes evident.

I do not claim, however, that the argument successfully targets a position according to which Sam is a substance cause of his action, while such substance-causation does not reduce to causation by events, but is nevertheless constituted by or grounded in causation solely by events (cf. Clarke 2017). Such a position endorses the metaphysical thesis that there is unreduced causation by agents fundamentally as substances, and thus denies that all causation is event-causation, a provision essential to the views the disappearing agent argument is designed to target. Thus by 'causation fundamentally as a substance' I do not mean to exclude substance causation wholly grounded in or constituted by event-causation, as long as it does not reduce to event-causation.⁵

5. The disappearing agent argument for agency

Mele (2015), Ekstrom (2016), and Clarke (2017) ask whether the disappearing agent objection extends to a successful objection to event-causal theories of action generally, whether indeterministic or deterministic. Mele remarks that given my rejection of agent-causal libertarianism, I'd be left without action, which is bad. But I'm open to the claim that the argument does so generalize (Pereboom 2015).

A number of philosophers argue that agency is intuitively agent-causal rather than event- or state-causal. In service of this contention, David Velleman nicely describes how action appears on the Davidsonian event-causal picture:

There is something the agent wants, and there is an action that he believes conducive to its attainment. His desire for the end, and his belief in the action as a means, justify taking the action, and they jointly cause an intention to take it, which in turn causes the corresponding movements of the agent's body. Provided that

⁴ Near the end of his comment, Mele outlines a position according to which an agent's moral responsibility on the event-causal view can result from the shaping of her values by past free decisions. I develop an objection to this type of view in (2001: 49).

⁵ I'm undecided about whether such a view falls to the disappearing agent objection. For my views on constitution, grounding, and reduction, see Chapter 7 of (Pereboom 2011). Thanks to Randy Clarke for prodding me to clarify my position on this issue.

these causal processes take their normal course, the agent's movements consummate an action, and his motivating desire and belief constitute his reasons for acting (Velleman 1992, 461).

Does this picture conflict with our intuitive sense of agency? John Bishop writes:

"Intuitively, we think of agents as carrying out their intentions or acting in accord with their practical reasons, and this seems different from (simply) being caused to behave by those intentions or reasons" (1989, 72). Velleman allows that certain non-paradigmatic cases of action, such as weak-willed action, fit the event- or state-causal picture, but that *full-blooded action* doesn't seem to:

In full-blooded action, an intention is formed by the agent himself, not by his reasons for acting. Reasons affect his intention by influencing him to form it, but they thus affect his intention by affecting him first. And the agent then moves his limbs in execution of his intention: his intention doesn't move his limbs by itself. The agent thus has at least two roles to play: he forms an intention under the influence of reasons for acting, and he produces behavior pursuant to that intention (Velleman 1992, 462).

Terry Horgan (2007) and Martine Nida-Rümelin (2007) explicitly invoke the phenomenology of agency in support of a view like Velleman's. Suppose you raise your arm and clench your fist. Horgan remarks:

You experience your arm, hand, and fingers as being moved by you yourself – rather than as experiencing their motion either as fortuitously moving just as you want them to move, or passively experiencing them as being caused by your own mental states. You experience the bodily motion as generated by yourself" (Horgan 2007, 187).

Nida-Rümelin adduces an additional claim about the phenomenology that would count against the event- or state-causal view. On her account, when one actively does something over moderately extended periods of time, such as slowly raising one's arm, the phenomenology is not as if one is experiencing a causal connection between one's intention and a bodily change (2007, 259). Horgan concurs; in the case of raising one's arm and clenching one's fist, it's not as if the phenomenology involves "first experiencing an occurrent wish for your right hand to rise and your fingers to move into clenched position, and then passively experiencing your hand and fingers moving in just this way," or first experiencing this occurrent wish "and then passively experiencing a causal process consisting of this wish's causing your hand to rise and your fingers to move into clenched position" (Horgan 2007, 186). These considerations give rise to a disappearing agent objection to event- or state- causal theories of action (Hornsby 2004a and b; Steward 2012)

Velleman (1992) proposes to solve this disappearing agent problem for agency within the event-causal framework by providing an account of the key role of the agent in event- or state-causal terms. His strategy is to account for the distinctive role of the agent in action by certain core desires or standing preferences, with which the agent, in its role in acting, can plausibly be identified. The role of the agent is played, specifically, by *a desire to act in accord with the best reasons* (Velleman 1992, 478- 79; 2000, 141; cf. Franklin 2014). This attitude is sufficient to constitute the role the agent has in acting:

Although the agent must possess an identity apart from the substantive motives competing for influence over his behavior, he needn't possess an identity apart from the attitude that animates the activity of judging such competitions. If there is such an attitude, then its contribution to the competition's outcome can qualify as his – not because he identifies with but rather because it is functionally identical to him (Velleman 1992, 480; Velleman 2000, 141)

The objection to this account that I want to raise is that it cannot account for the settling role in action that the agent intuitively can have in situations of rational equipoise. Ralph is torn about whether to stay in Mayberry or move to New York, and his reasons for each option are and appear to him to be in equipoise. Yet he just chooses to move to New York. It can't be a desire to act in accord with the best reasons that settles how Ralph acts, since this desire does not favor moving to New York over staying in Mayberry. Thus given Velleman's functional account of the agent, Ralph can't settle which option becomes actual, and so in this sort of situation he can't act. This counts against Velleman's proposal (Pereboom 2015; cf. Runyan 2016, 1633-34; this objection also challenges against Franklin's (2014) development of Velleman's view).

Laura Ekstrom (2000, 2003, 2016) develops a related account, one that ties the agent to her undefeated authorized general preferences, where such preferences are non-coercively formed or maintained, and are caused but not causally determined by considerations brought to bear in his deliberation. But this account is also subject to the objection that it can't account for settling in equipoise situation. In Robert Kane's (1996) example of a torn decision, a businesswoman can either decide to stop and help an assault victim for moral reasons, or else decide to speed on to work for self-interested reasons, and these reasons are in motivational equipoise. We can imagine that she has both moral and self-interested undefeated authorized general preferences. What if these general preferences are in motivational equipoise as well, so that now the decision would not only be torn, but in this sense meta-torn? Still, intuitively, the businesswoman can still settle whether to stop and help or speed on to work. This counts against Ekstrom's proposal.

6. Deterministic agent causation

If agent causation is required to solve the disappearing agent problem for event-causal libertarianism, now it appears to be required to solve a disappearing agent problem that arises for the event-causal theory of agency more generally. However, I pose a problem for agent-causal libertarianism that motivates me to accept free will skepticism instead (Pereboom 2001, 2014). This leads Mele (2015) to ask whether I am now left without an account of agency. But at this point it's open to me to invoke a notion of agent causation compatible with the causal determination of action (Pereboom 2015), since, as I explain below, the difficulty I raise for agent-causal libertarianism does not extend to deterministic agent causation, in recent times proposed by Ned Markosian (1999, 2010) and Dana Nelkin (2011). As on the libertarian version, agency (or at least full-blooded agency) is accounted

for by the existence of agents who as substances have the power to cause actions. But by contrast with the libertarian counterpart, in the exercise of their agent-causal power agents are in general causally determined by factors beyond their control.

The difficulty I raise for agent-causal libertarianism is that it cannot be reconciled with what we would expect given our best physical theories (Pereboom 1995, 2001, 2014). Consider first the supposition that all the events in the physical world are governed by deterministic laws. In the agent-causal libertarian picture, when an agent makes a free decision she causes the decision without being causally determined to do so. On the path to action that results from this undetermined decision, changes in the physical world, for instance in her brain or some other part of her body, are produced. But it would appear at this point we would encounter divergences from these laws. The changes in the physical world that result from the undetermined decision would themselves not be causally determined, and they would thus not be governed by deterministic laws. One might object that it is possible that the physical changes that result from every free decision just happen to dovetail with what could in principle be predicted on the basis of the deterministic laws, so nothing actually occurs that diverges from these laws. But this proposal would seem to involve coincidences too wild to be credible. Thus agent-causal libertarianism is not plausibly reconcilable with the physical world's being governed by deterministic laws.

On some interpretations of quantum mechanics, the physical world is not in fact deterministic, but is rather governed by probabilistic laws. But wild coincidences would also arise on this suggestion. Consider the class of possible actions each of which has a physical component whose antecedent probability of occurring is approximately 0.32. It would not violate the laws in the sense of being logically incompatible with them if, for a large number of instances, the physical components in this class were not actually realized close to 32% of the time. Rather, the force of the statistical law is that for a large number of instances it is correct to expect physical components in this class to be realized close to 32% of the time. Are free choices on the agent-causal libertarian model compatible with what the probabilistic law would have us to expect about them? If they were, then for a large enough number of instances the possible actions in our class would almost certainly be freely chosen near to 32% of the time. However, if the occurrence of these physical components were settled by the choices of agent causes, then actually being chosen close to 32% of the time would amount to a wild coincidence. The proposal that agent-caused free choices do not diverge from what the statistical laws predict for the physical components of our actions would be so sharply opposed to what we would expect as to make it incredible (Pereboom 1995, 2001).

But deterministic agent causation is not undermined by this concern about reconciliation with the laws of nature. If the laws of nature are deterministic, and agent causation is deterministic, then instances of agent causation might all be governed by the laws of nature.

I don't believe, however, that the disappearing agent argument that aims at event-causal theories of action generally is as strong as the one that targets only event-causal libertarianism. This is because the more general argument requires an equipoise situation, while the less general one does not. Event-causal libertarians agree that agents have freedom of choice in non-equipoise situations, and the settling problem arises in such cases as well. But

event-causal determinists can hold that situations in which motivating reasons don't causally determine scarcely ever arise, if at all, and if they do, the agent can deploy a psychological device functionally similar to the throw of the dice, instead of settling by just choosing. Thus event-causal determinism remains a viable option.⁶

7. The present luck and enhanced control objections

One objection to event causal libertarianism that both Clarke and I have voiced (Pereboom 2001, 2007; Clarke 2003) is the “no enhanced control” argument. The idea is that if it's agreed that agents whose actions are causally determined by factors beyond their causal reach lack the control in acting required for moral responsibility, mere addition of indeterminacy won't supply the requisite enhanced control. I said:

If factors beyond the agent's control, rather than determining a single decision, instead simply leave open which decision will occur, and the agent has no greater role in the production of the decision than she does in the deterministic context, then there is no more reason to think that she is morally responsible than there is in the deterministic context. So it appears that the event-causal libertarian can supplement the deterministic context only with the relaxation of the causal net. (Pereboom 2007, 195)

But Christopher Franklin (2016) develops a response to this argument, one that persuades me. Given that I allow for deterministic agent causation, and that I am open to the possibility that indeterministic agent causation does secure the sort of free will required for moral responsibility, then relative to deterministic agent causation mere relaxing of the causal net may indeed precipitate the sort of control required for moral responsibility. So I cannot accept the general principle that mere relaxation of the causal net cannot enhance control in the relevant way, a principle essential to the “no enhanced control” objection.

Consider next the present luck objection to event-causal libertarianism (Mele 1999, 2006; Haji 2004; Clarke 2017). Suppose that *S's deciding to A at t1* satisfies the event-causal libertarian's conditions for free action. Then, given exactly the same conditions antecedent to *t1* as those that precede *S's* decision in the actual world, *S's deciding to A at t1* might not have occurred. So there is a non-actual possible world, *W**, that shares a history up to *t* with the actual world, *W*, but while *S's deciding to A at t1* occurs in *W*, it fails to occur in *W**. Thus the fact that *S's deciding to A at t1* did actually come about would seem to be a matter of responsibility-undermining luck.

But I argue that present luck conditions alone are insufficient to undermine moral responsibility (Pereboom 2001, 2014; cf. Griffith 2010). If *S* agent-causes *A* at *t1* (without being causally determined to do so by factors beyond her control), an event of the following type occurs:

G: *S's* agent-causing *A* at *t*.

⁶ Thanks to Colin McLear and Simon-Pierre Chevarie-Cossette for valuable discussion of this issue.

As Mele (1999, 2006) and Haji (2004) emphasize, given the same history antecedent to t_1 as those that precede G , and given the indeterminism of the agent-causal libertarian view, G might not have occurred. Thus there is a non-actual possible world, W^* , that shares a history up to t_1 with the actual world, W , but while G occurs in W , it fails to occur in W^* . But is the fact that G occurred merely a matter of responsibility-undermining luck? On my preferred version of the agent-causal libertarian view, what the agent-as-substance does *most fundamentally* is to cause A , and the proposal is that it is in causing A that her responsibility-conferring control over A is located. The substance-causal relation is embedded in event G , and thus G will not be what is most fundamentally caused. The agent-causal libertarian maintains that the crucial control is not exercised by way of events that precede G , but by the agent-as-substance. If in addition to the events that precede G we hold fixed in W and W^* the agent-as-substance's exercise of her agent-causal power, G will occur in W and not in W^* , but only because the agent-as-substance causes the action A in W but not in W^* . For this reason, it isn't a matter of responsibility-undermining luck that G occurs in W .

This last claim is indeed controversial (Clarke 2010), but I find it compelling. Thus in my view, present luck does not, in general, preclude the control required for moral responsibility. Note further that for both the "no enhanced control" argument and the present luck argument, the cases in which the consideration adduced fails to preclude responsibility are those that feature an agent who has the power to settle which action occurs. In the deterministic agent-causal context, relaxing the causal net does yield enhanced control. In the indeterministic context with agent causation, variation in action with the same causal history up to the time of the action does not yield responsibility-undermining luck. This indicates that it's the disappearing agent argument that focuses our attention on the real problem for the event-causal libertarian view.⁷

⁷ Thanks to audiences at the Collège de France in September 2016 and at Balliol College, Oxford University in May 2017 for enlightening discussions. Thanks also to Simon-Pierre Chevarie-Cossette, Randy Clarke, Laura Ekstrom, Chris Franklin, Richard Holton, Colin McLearn, Michael McKenna, Al Mele, Dana Nelkin, Tim O'Connor, and Carolina Sartorio for valuable comments and interaction.

References

- Balaguer, M. 2010. *Free Will as an Open Scientific Problem*, Cambridge MA: MIT Press.
- Balaguer, M. 2014 "Replies to McKenna, Pereboom, and Kane," *Philosophical Studies* 169, pp. 71-93.
- Bergson, H.. 1889/1910. *Essai sur les données immédiates de la conscience*, Paris: F. Alcan, 1889; translated as *Time and Free Will* by F. L. Pogson, London: Allen and Unwin, 1910.
- Bishop, J. 1989. *Natural Agency*, Cambridge: Cambridge University Press.
- Chisholm, R. 1971. "Reflections on Human Agency," *Idealistic Studies* 1, pp. 33-36.
- Clarke, R. 2003. *Libertarian Accounts of Free Will*, New York: Oxford University Press.
- Clarke, R. 2010. "Are We Free to Obey the Laws?" *American Philosophical Quarterly* 47, pp. 389-401.
- Clarke, R. 2013. "Some Theses on Desert," *Philosophical Explorations* 16, pp. 153-64.
- Clarke, R. 2017. "Free Will, Causation, and the Disappearing Agent Argument," *Noûs*, DOI 10.1111/nous.12206
- DeRose, K. 1993. "Review of William Rowe's *Thomas Reid on Freedom and Morality*," *Philosophy and Phenomenological Research* 53, pp. 945-49.
- Ekstrom, L. W. 2000. *Free Will, A Philosophical Study*, Boulder CO: Westview.
- Ekstrom, L. W. 2003. "Free Will, Chance, and Mystery," *Philosophical Studies* 113, pp. 153-180.
- Ekstrom, L. W. 2016. "Towards a Plausible Event-Causal Account of Free Will," *Synthèse* DOI 10.1007/s11229-016-1143-8.
- Feinberg, J. 1970. *Doing and Deserving*, Princeton, NJ: Princeton University Press.
- Franklin, C. 2011. "Farewell to the Luck (and Mind) Arguments," *Philosophical Studies* 156, pp. 199-230.
- Franklin, C. 2014. "Event-Causal Libertarianism, Functional Reduction, and the Disappearing Agent Argument," *Philosophical Studies* 170, pp. 413-32.
- Franklin, C. 2016. "If Anyone Should be an Agent-Causalist, then Everyone Should be an Agent-Causalist," *Mind* 125, pp. 1101-1131.
- Ginet, C. 1997. "Freedom, Responsibility, and Agency," *Journal of Ethics* 1, pp. 85-98.
- Ginet, C. 2007. "An Action Can Be Both Uncaused and Up to the Agent," in C. Lumer ed. *Intentionality, Deliberation, and Autonomy*, Ashgate, pp. 243-56.
- Goetz, S. 2008. *Freedom, Teleology, and Evil*, London: Continuum, 2008.
- Griffith, M. "Why Agent-Caused Actions Are Not Lucky," *American Philosophical Quarterly* 47, pp. 43-56.
- Haji, I. 2004. "Active Control, Agent Causation, and Free Action," *Philosophical Explorations* 7, pp. 131-48.
- Hornsby, J. 2004. "Agency and Actions," in *Agency and Action*, H. Steward and J. Hyman, eds. Cambridge: Cambridge University Press, pp. 1-23.
- Kane, R. 1996. *The Significance of Free Will*, Oxford: Oxford University Press.
- Lucretius (50 BCE). *De Rerum Natura*, tr. W.H.D. Rouse, revised by M. F. Smith, Cambridge: Harvard University Press.
- McCann, H. (1998). *The Works of Agency*, Ithaca: Cornell University Press.
- Mele, A. 1992. *Springs of Action*. New York: Oxford University Press.

- Mele, 1999. "Ultimate Responsibility and Dumb Luck," *Social Philosophy & Policy*, 16, pp. 274–93.
- Mele, A. 2006. *Free Will and Luck*. New York, Oxford University Press.
- Mele, A. 2015. "On Pereboom's Disappearing Agent Argument," *Criminal Law and Philosophy*. DOI: 10.1007/s11572-015-9374-1.
- Markosian, N. (1999). "A Compatibilist View of the Theory of Agent Causation," *Pacific Philosophical Quarterly* 80 (1999), pp. 257-77.
- Markosian, N. 2010. "Agent Causation as the Solution to All the Compatibilist's Problems," *Philosophical Studies* 157, pp. 383–98.
- Nelkin, D. 2011. *Making Sense of Freedom and Responsibility*, Oxford: Oxford University Press.
- Nida-Rümelin, M. 2007. "Doings and Subject Causation," *Erkenntnis* 67, pp. 255-72.
- O'Connor, T. 2000. *Persons and Causes*, New York: Oxford University Press.
- O'Connor, T. 2008. "Agent-Causal Power," in *Dispositions and Causes*, ed. Toby Handfield, Oxford: Oxford University Press, pp. 189-214.
- Pereboom, D. 1995. "Determinism *Al Dente*," *Noûs* 29, 1995, pp. 21-45.
- Pereboom, D. 2001. *Living without Free Will*, Cambridge: Cambridge University Press.
- Pereboom, D. 2004. "Is Our Conception of Agent-Causation Coherent?" *Philosophical Topics* 32, pp. 275-86.
- Pereboom, D. 2011. *Consciousness and the Prospects of Physicalism*, New York: Oxford University Press.
- 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 *Horizons of Authenticity in Phenomenology, Existentialism, and Moral Psychology*, H. Pedersen and M. Altman, eds., New York: Springer, pp. 277-94.
- Pereboom, D. 2017. "A Defense of Free Will Skepticism: Replies to Victor Tadros, Saul Smilansky, Michael McKenna, and Alfred Mele on my *Free Will, Agency, and Meaning in Life*," *Criminal Law and Philosophy*, DOI:10.1007/s11572-017-9412-2.
- Runyan, J. D. 2016. "Events, Agents, and Settling Whether and How One Intervenes," *Philosophical Studies* 173, pp. 1629-1646.
- Sartorio, C. 2013. "Making a Difference in a Deterministic World," *Philosophical Review* 122, pp. 189-214.
- Sartorio, C. 2016. *Causation and Free Will*, Oxford: Oxford University Press.
- Steward, H. 2012. *A Metaphysics for Freedom*, Oxford: Oxford University Press.
- Velleman, J. D. 1992. "What Happens When Someone Acts?" *Mind* 101, pp. 461-481.
- Velleman, J. D. 2000. "Introduction," in J. D. Velleman, ed., *The Possibility of Practical Reason*, New York: Oxford University Press, pp. 1-30.