

## Is Agent Causation Possible?

Noah McKay  
*University of Edinburgh*

**ABSTRACT:** To meet the luck objection to incompatibilism, philosophers such as Timothy O'Connor, Randolph Clark, and William Rowe resurrected the Reidian notion of agent causation, which implies the "Substance-Causal Thesis" (SCT): some causes are fundamentally substances, not events. I examine an objection to SCT by C. D. Broad, developed by Carl Ginet, that substances cannot cause events because substances cannot explain why events happen when they do. The objection fails as it rests on a demand for contrastive explanations of free actions. However, I will show that a slightly different objection succeeds in showing SCT false.

In recent years, incompatibilists about free will have faced a barrage of "luck" objections to their various accounts of free action. These objections, in broad form, go like this: If some free action,  $A_1$ , is caused indeterministically by some cluster of mental events,  $E$ , which might equally well have caused an alternative action,  $A_2$ , then it seems to be nothing more than a matter of luck that  $A_1$  occurs rather than  $A_2$ . At the very least, it does not seem right to say that it is up to anyone whether  $A_1$  or  $A_2$  occurs, any more than it would be if  $E$  caused  $A_1$  deterministically. Put another way, incompatibilists simply inject an element of chance into free action that is absent on the determinist picture, and it is hard to see how this element of chance gives an agent any more control over his actions than he would have otherwise.<sup>1</sup>

The way out of the luck objection, according to most incompatibilists, is to offer an independent account of how agents in an indeterministic world can exercise a robust kind of control over their actions. In pursuit of such an account, several authors have proposed a return to the long out-of-fashion notion of agent causation, famously defended by the 18<sup>th</sup>-century Scottish philosopher Thomas Reid. In his *Essays on the Active Powers of Man*, Reid argues that free agents, rather than their motives, are the efficient causes of their free actions (202–6, 215–7). In more contemporary language, when an agent performs a free action, the agent himself, considered as a substance, directly causes

the action rather than any mental event or events of which he is the subject.

William Rowe, Timothy O'Connor, and Randolph Clarke have recently developed and defended theories of agency similar to (and, in Rowe's and O'Connor's cases, explicitly derivative of) Reid's. These authors hope that agent causation will supply the control over free actions that the incompatibilist view needs to overcome the luck objection. As Clarke puts it:

On these agent-causal accounts [...] an agent is in a strict and literal sense an originator, an initiator, an ultimate source of her directly free action; she is an uncaused cause of that behavior, and one whose causing of that behavior is not causally determined. The required indeterminism and origination is thought to secure both the genuine openness of alternatives and the exercise of active control that figure in the freedom that we desire. [134]

Some philosophers have protested that the agent-causation view is still vulnerable to the luck objection or a modified version of it (substitute "agent" for "cluster of mental events" in the opening paragraph, for instance). However, in this essay, I consider a more fundamental objection: that agent causation is impossible.

Despite the considerable variety among agent-causal theories, they all imply what I call the "Substance-Causal Thesis":

(SCT) Some causes are fundamentally substances, not events.<sup>2</sup>

SCT is a necessary condition of the claim, common to all agent-causalists, that, in O'Connor's words, "[In] instances of *agent causation*, the cause of an event is not a state of, or event within, an agent; rather, it is the agent himself, an enduring substance" ("Agent-Causal Power," 193). In this essay, I examine an influential challenge to the possibility of SCT first put forth by C. D. Broad, further developed by Carl Ginet, and addressed by Clarke and O'Connor. I then offer an original objection to the possibility of SCT that draws on premises similar to those of Broad's argument but that is immune to the objections O'Connor and Clarke level against it.

SCT may at first appear unintelligible. What could it mean that a substance, *qua* substance, causes an event to occur? As O'Connor points out, while it is common in everyday speech to describe substances as causes, this is clearly a shorthand way of saying that events *involving* those substances cause things to happen. While we might say that a car knocked down a telephone pole, what we really mean is that the movement of the car toward the telephone pole, or the collision of the car with the telephone pole, caused it to fall. And these causes are events (O'Connor, "Agent as Cause"). It is difficult to grasp how a substance, as *distinct* from any event in which it participates, could cause a telephone pole to fall.

Before I continue, I should be a bit (but not much) clearer about what I mean by "event." There is much disagreement among metaphysicians about precisely what events are, but there is a consensus that where there are property-instances, there are events, and vice versa. That is, when anything bears a property or comes to bear a new property (or stands in a relation or comes to stand in a new relation), an event occurs. So, the sentence "The light turned green" describes an event and likewise for the sentences "The light was red" and "The light stood at a distance from me." This implies that any claim that such-and-such an object bears such-and-

such a property implicates that object in an event. (This will be important later.)

C. D. Broad was one of the first to give argumentative teeth to the mysteriousness of agent causation:

The putting-forth of an effort of a certain duration, is quite clearly an event or process, however unique and peculiar it may be in other respects. It is therefore subject to any conditions which self-evidently apply to every event, as such. Now it is surely quite evident that, if the beginning of a certain process at a certain time is determined at all, its total cause *must* contain as an essential factor another event or process which *enters into* the moment from which the determined event or process *issues* [...]. How could an event possibly be determined to happen at a certain date if its total cause contained no factor to which the notion of date has any application? And how can the notion of date have any application to anything that is not an event? [215]

Broad seems to suggest that, unless the cause of an event occurs at a time, there is no explanation for why that event occurs precisely when it does. And since agents *qua* substances do not occur at times, they cannot explain why free actions occur just when they do. Carl Ginet offers a somewhat clearer statement of the problem:

On the agent-causal theory, the immediate cause of the occurrence of a particular sort of simple mental event at a particular time is the agent herself, *per se* and not in virtue of any event of which she is the subject. But the agent *per se* cannot explain why the event happened precisely when it did rather than at some slightly different time. Only some difference between the agent at the one time and the agent at the other times, some temporally located property, could do that. Nor, it might be added, can the agent *per se* explain why that particular sort of event rather than some other sort happened just then. What sense can it make, then, to say that the agent as

such is the cause of the occurrence of that particular sort of event rather than some other sort, and is the cause of its occurring at that particular time rather than some other time? [94]

At first blush, this appears to be a rather grave difficulty. But the defenders of agent causation have offered rejoinders.

Clarke claims that the agent-causalist can solve this problem by adopting an *integrated* account of agent causation according to which free actions are caused by agents *qua* substances *and* (indeterministically) by certain mental events. On Clarke's view, it is a matter of natural law that an agent directly causes a free action when and only when that agent's reasons indeterministically cause that free action in the same way that, as a matter of natural law, a spark causes a fire when and only when the presence of fuel and oxygen in its vicinity also causes that fire (144–8). (Notice that, while Clarke's view incorporates event-causation of free actions, he is still committed to SCT.) By including the agent's reasons for an action in the causal explanation of that action, Clarke argues, we may explain the timing of the action in terms of the timing of those reasons (198–9).

This solution, of course, depends on the plausibility of the claim that, as a matter of natural law, free actions are caused by reasons when and only when they are caused by agents. O'Connor objects that this view pits agents against their reasons for the role of causal progenitor of free actions (*Persons and Causes*, 78). But I do not think it is necessary to interpret Clarke this way: oxygen, fuel, and heat do not *compete* for causal roles in the production of a fire. Rather, they *jointly* cause a fire. And I see no reason why Clarke cannot maintain that agents cause free actions jointly with their reasons in a similar way.

Michael McKenna and Derek Pereboom object to Clarke's account on grounds that it requires the positing of a brute law connecting agent-causation with event-causation for "there would appear to be nothing about the agent-as-substance *per se* that explains why its propensities to cause actions match those of the reasons"

(247). This objection, it seems to me, is on-target: If agent-causal power and reasons are held to be joint, independent causes of free actions in the sense that oxygen, fuel, and heat are joint, independent causes of combustion, it is very odd that they are so perfectly coordinated, in the same way that it would be odd if fuel, oxygen, and heat always turned up together, and never separately.<sup>3</sup>

O'Connor rejects Ginet's challenge on different grounds. According to O'Connor, Ginet's objection depends on an unreasonable demand for *contrastive* explanations of actions—explanations of why some actions are performed *rather than* others or performed at particular times *rather than* others. (By contrast—no pun intended—a noncontrastive causal explanation of an action, A, would simply cite the causal factors that led to A without citing those factors to explain why A occurred *rather than* any possible alternative.) As O'Connor points out, any indeterminist will be hard-pressed to provide contrastive explanations for a whole bevy of facts (91–5). Even those incompatibilists who reject agent causation claim that free actions are the causal products of mental events that might just as well have caused different actions from those which they in fact caused. So even a run-of-the-mill libertarian like Ginet will, on any particular occasion, lack a contrastive causal explanation for why someone performed one particular kind of free action rather than another or why she performed it just when she did rather than at some other time. The most he can do is provide a noncontrastive explanation of the action in question. And that is no more than the agent-causalist can do.

In my view, this is a successful reply. Insofar as it is plausible for incompatibilists to maintain that free actions are indeterministically caused, it is plausible to deny that free actions have the kind of contrastive explanations that Broad and Ginet are after. However, I think that their objection can be recast in a way that preserves the intuition that agents *qua* substances lack an important kind of explanatory adequacy while eschewing any demands for contrastive explanations.

In the course of developing Broad's objection, Ginet makes the following observation:

Broad was, I believe, assuming that it is incoherent to suppose that the cause of something's being the case might fail to explain its being the case, and I am inclined to agree. If X causes it to be the case that Y rather than any alternative then there must be something about X that explains why Y rather than any alternative. [95]

This claim is highly plausible, especially if it is weakened so that (i) in order for X to be a cause of Y, X need only noncontrastively explain Y, and (ii) in order for X to be a cause of Y, X need only partially explain Y. (Consider what would follow if this were false: there would exist some X and some Y such that X caused Y but X did not even partially explain Y. Clearly, that is not possible; if anyone maintained that it was, I would have to infer, in Reidian fashion, that he did not understand the word "cause.") This premise, I contend, may function as the basis for a more successful objection to SCT. Although Broad and Ginet both restrict their discussions to the inadequacy of agents *qua* substances to provide contrastive explanations of actions, there is a case to be made that agents *qua* substances are inadequate to provide any kind of explanation at all.

SCT posits that some substances function as causes independently of any events in which they participate. If a cause of an event cannot fail to explain that event, then it follows that substance-causes must explain their effects *independently* of any events in which they participate. As noted earlier, where there are property-instances, there are events, so any explanation that cites property-instances necessarily cites events. But this means that substance-causes must explain their effects independently of any properties they exemplify (or relations they stand in). That is, there must be some substance, S, and some event, E, such that S explains E but *nothing at all about* S explains E. For if something about S explained E, then the true explanation of

E would cite a property or relation exemplified by S. And this is sufficient for citing an event. Hence, S would not explain E independently of any events in which it participated and would therefore fail to be a substance-cause after the manner of SCT. But clearly, it is absurd to say that S, but nothing at all about S, explains E.

I hardly think that this last claim needs defending, but here is an argument for the exceptionally stubborn: If S explains E, but nothing about S explains E, then some proposition about S but not about any of S's properties explains some proposition about E. There is only one remotely plausible candidate for a proposition that is about S but not about any of S's properties: *S exists*. But clearly, the mere fact that an agent exists does not in any way explain his actions.

Hence, SCT is false. (Indeed, if we take Ginet's premise above to be a metaphysically necessary truth, then SCT is strictly impossible.) This objection does not rest upon any demand for contrastive explanations of actions, so O'Connor's rejoinder to Ginet will not resolve it. And Clarke cannot invoke an integrated agent-causal account to handle the problem because whether an agent's reasons at the time of her action explain the timing or nature of that action is entirely irrelevant to whether the *agent* meets the conceptual requirements to be a cause of that action. (Even if the agent's reasons explain the action, and even if the agent features in a reasons-explanation of the action, it does not follow that the agent *qua* substance explains the action. It follows only that the agent *qua* reasoning subject explains it.)

I do not pretend to know how O'Connor would attempt to meet this challenge. While addressing a related objection, Clarke however offers a reply that could be advanced against my argument:

Explanations that answer why-questions, however, may not be the only causal explanations. Sometimes an effect of a certain type can be brought about in a number of different ways; where that is so, we may want to know how it was brought about in this case.

A causal explanation of this sort answers what we may call a “how-actually” question. And it may be claimed that citing the agent in a case where an action is performed freely, as characterized by an integrated agent-causal account, tells us how something that could have been brought about without substance causation was, in fact, brought about in this case. [*Libertarian Accounts*, 199–200]

The core of Clarke’s reply seems to be this: Even if agents *qua* substances do not explain *why* free actions occur, they explain *how* actions occur. And that is good enough.

I do not think this response is satisfactory. The claim that an agent *qua* substance explains *how* a free action occurred is conceptually indistinct from the claim that the agent *qua* substance was, in fact, the cause of that free action. (This is implied in the passage quoted above: for Clarke, a how-explanation is just a specification of which of a number of potential factors caused an event. So, an agent-caus-

al how-explanation merely specifies the agent *qua* substance as the cause of a particular event.) And this is the very claim at issue. Alternatively, we could restate the critical premise of the objection thus: If X explains how Y occurred, then X explains why Y occurred. This modified premise is no less plausible than the original.

In this essay, I have examined an objection to the possibility of agent causation first presented by C. D. Broad and developed by Carl Ginet. I have concluded that the objection in its original form fails since it rests on an unreasonable demand for contrastive explanations of free actions. I have presented, however, a modified version of the objection that avoids this dependency. In summary, I have argued that agents *qua* substances cannot cause free actions because agents *qua* substances, divorced as they are from any of their properties, are inadequate to explain anything. If I am right, SCT, and hence agent-causation, is impossible. Incompatibilists will need an alternative strategy to combat the luck objection.

### Notes

<sup>1</sup> For a helpful survey of luck objections, see Pereboom and McKenna, 236–9.

<sup>2</sup> “Fundamentally” is an important qualifier here. Some philosophers, such as Roderick Chisholm (in his later work), are considered to be agent-causal libertarians despite their denial that agent causation is fundamentally a case of substance causation. When I speak of “agent causation,” I mean to exclude those theories of action that take agency to be ultimately ontologically reducible to event-causation.

<sup>3</sup> This particular example of joint causation is Clarke’s.

### Works Cited

- Broad, C. D. “Determinism, Indeterminism, and Libertarianism.” In *Ethics and the History of Philosophy*, edited by C. D. Broad, 195–217. London: Routledge, 1952.
- Clarke, Randolph. *Libertarian Accounts of Free Will*. Oxford: Oxford University Press, 2003.
- Carl Ginet, Carl. “Freedom, Responsibility, and Agency.” *The Journal of Ethics* 1, no. 1 (1997): 85–98.
- O’Connor, Timothy. “The Agent as Cause.” In *Metaphysics: The Big Questions*, edited by Dean W. Zimmerman and Peter van Inwagen, 374–80. Oxford: Blackwell, 2008.
- . “Agent-Causal Power.” In *Dispositions and Causes*, edited by Toby Handfield, 189–214. Oxford: Clarendon Press, 2009.
- O’Connor, Timothy. *Persons and Causes: The Metaphysics of Free Will*. New York: Oxford University Press, 2000.
- Pereboom, Derek and Michael McKenna. *Free Will: A Contemporary Introduction*. New York: Routledge, 2016.
- Reid, Thomas. *Essays on the Active Powers of Man*, edited by Knud Haakonssen and James A. Harris. Edinburgh: Edinburgh University Press, 2010.
- Rowe, William. *Thomas Reid on Freedom and Morality*. London: Cornell University Press, 1991.