

FREEDOM, RESPONSIBILITY, AND AGENCY

(Received and accepted 18 September 1995)

ABSTRACT. This paper first distinguishes three alternative views that adherents to both incompatibilism and PAP may take as to what constitutes an agent's determining or controlling her action (if it's not the action's being deterministically caused by antecedent events): the indeterministic-causation view, the agent-causation view, and "simple indeterminism." The bulk of the paper focusses on the dispute between simple indeterminism – the view that the occurrence of a simple mental event is determined by its subject if it possesses the "actish" phenomenal quality and is undetermined by antecedent events – and Timothy O'Connor's agent-causation view. It defends simple indeterminism against O'Connor's objections to it and offers objections to O'Connor's view.

KEY WORDS: agency, agent-causation, control, determinism, free will, incompatibilism, indeterministic causation, Principle of Alternative Possibilities, responsibility

Let us say that an action is free if and only if up until the time of the action the agent had it open to her not to perform it: she could then have performed some other action instead or not acted at all. I mean this to be a stipulative definition of free action, not a substantive thesis about it.

There are, however, two substantive, controversial theses about free action, thus defined, that I am going to assume for the purposes of this paper. The problem I want to discuss here arises only for someone (like me) who holds both of these theses. One is the thesis that freedom of action is incompatible with determinism, that free action cannot occur in a completely deterministic universe, a universe where the laws of nature and the state of the world at any given time determine everything that happens after that time. This thesis is usually referred to as incompatibilism.¹ The other thesis is that an agent can be morally responsible for her action only if it is a free action: an agent can merit credit or blame for something she did only if she could have done otherwise. This thesis is commonly referred to as the Principle of Alternative Possibilities, or PAP for short.²

¹ I have argued for this thesis in Ginet, *On Action* (Cambridge: Cambridge University Press, 1990), Chap. 5.

² I defend this thesis in Ginet, "In Defense of the Principle of Alternative Possibilities: Why I Don't Find Frankfurt's Argument Convincing," forthcoming in J. Tomberlin (ed.), *Philosophical Perspectives 10: Metaphysics* (Atascadero, CA: Ridgeview, 1996).

It follows from the conjunction of incompatibilism and PAP that an agent is responsible for her action only if that action was not causally determined by antecedent states and events. This consequence may seem – has seemed to many – to present a serious problem. If an action was not causally determined by any antecedent states and events, not even ones in the agent, then how can it be right to say that the agent exercised control over its occurrence, *determined whether* it or something else would occur; and if she did not do that, how can she be responsible for its occurrence? Surely an agent is responsible for her action – or for any other occurrence – only if she made it happen; but how can she, or anything, have made it happen if nothing caused it? But if PAP is right, the agent can be responsible for her action only if she was free to do otherwise, and, if incompatibilism is right, she was free only if her action was not causally necessitated by antecedent events. So it seems that, given the reasonable-sounding principle that an agent is responsible for her action only if she determines whether it happens, the conjunction of PAP and incompatibilism has the consequence that an agent is morally responsible for an action only if the action was both caused and uncaused.

We PAP incompatibilists seem to be stuck with an incoherent account of moral responsibility. We can provide a satisfactory metaphysical foundation for moral agency only if we can dispel that appearance. I believe we can, but at the moment we are not in agreement on how to do it. In what follows I try to explain this disagreement and to defend the way I think is best.

Timothy O'Connor examines three different sorts of response to the problem.³ The reasoning that raises the problem (the reasoning from the conjunction of PAP and incompatibilism to the unpalatable conclusion that our concept of moral agency is incoherent) relies on an additional premise, namely, the claim that if an action was not determined (in the sense of deterministically caused) by antecedent states and events, then it cannot be that the agent was responsible for it because it cannot be that the agent controlled or determined its occurrence. All three responses deny this additional premise. They all say that they can explain how, despite the action's not being causally necessitated by antecedents, it can nevertheless be right to say that the agent determined or controlled its occurrence. But the three responses differ radically in the accounts they give of what can make it right to say this.

On one view, the agent can be said to have determined or brought about the action if the action has a causal explanation in terms of the reasons

³ See T. O'Connor's instructive paper, "Indeterminism and Free Agency: Three Recent Views," *Philosophy and Phenomenological Research* 53 (1993), pp. 499–526.

for which the agent performed the action. This view holds, as I believe any incompatibilist must hold, that an agent's doing an action for reasons is compatible with the action's not being deterministically caused by the motives that supply her reasons or by any other antecedents. And it holds that a reasons explanation is a kind of *indeterministic* causal explanation, depending on merely probabilistic laws, and it is the fact that an action has this sort of explanation that makes the action something that the *agent* determined. Let us refer to this as the *indeterministic-causation* view.

A second view – which is the view I favor – has it that any attempt to explain an agent's determining of her own action in terms of this or that special sort of cause of it is unnecessary. For an agent to determine whether or not an event, *e*, occurs is for her to make it the case that *e* occurs by performing some suitable free action. If *e* is *not* her own free action then causation must enter into what it is for her to make it the case that *e* occurs: she can do this only by performing some free action that causes *e*. But if *e* is her own free action, then she makes it the case that *e* occurs, not by causing it, but by simply performing it (This latter “by” is logical rather than causal; we have a causal “by” in “I made a C major chord sound by pressing those three keys;” we have a logical “by” in “I made a C major chord sound by making sound simultaneously a C, an E, and a G”). Given that the action is free, the agent determines it, one could say, simply by being its subject, the one whose action it is. That is to say, all free actions are *ipso facto* determined by their subjects. O'Connor dubs this view *simple indeterminism*.

Third, and finally, there is the *agent-causation* view that O'Connor himself favors. This view affirms that the agent controls her action only if there is a direct causal relation between the agent and the action (or some event internal to the action), a causal relation where the relatum on the cause side is not any event in or state of the agent but just the agent herself, that enduring entity. The agent-causation view was put forward in the eighteenth-century by Thomas Reid⁴ and endorsed in the latter half of the present century by Roderick Chisholm⁵ and Richard Taylor.⁶ It has been subjected to criticism that many (including me) have thought fatal

⁴ T. Reid, *Essays on the Active Powers of the Human Mind* (Cambridge: MIT Press, 1969), Essay I.

⁵ R.M. Chisholm, “Freedom and Action,” in K. Lehrer (ed.), *Freedom and Determinism* (New York: Random House, 1966), pp. 11–44.

⁶ R. Taylor, *Action and Purpose* (Englewood Cliffs, NJ: Prentice-Hall, 1966), Chap. 9.

and has not been widely held in recent decades. But O'Connor has made a challenging attempt to revive it.⁷

O'Connor and I agree in rejecting the indeterministic-causation account of what it is for an agent to determine her action. O'Connor's dissatisfaction with this view arises from his belief that any such account must explain how an agent makes one among the competing motives she faces the efficacious one, and he doesn't see any way that this can be done without resorting to agent-causation. He considers two specific attempts to do it, one by Robert Nozick⁸ and one by Robert Kane,⁹ and provides illuminating criticism of both.

My reasons for finding the indeterministic-causation account unsatisfactory differ from O'Connor's. One difficulty I see is that there can be free and responsible actions that do not have reasons explanations. An agent can just spontaneously do something for no particular reason and not in order to satisfy any antecedent desire or intention she may have had (Of course, at the time of an action the agent must always intend to be doing it, or at least the basic action that is its initial phase. But this is a trivial truth about an action, pointing to an intrinsic feature of it that is entailed by its being an action. It cannot provide the sort of *cause* of the action that this account thinks we must have if the agent is to be responsible for her action). Another problem with the indeterministic causation account is that reasons explanations of actions do not require for their truth the truth of any causal laws (probabilistic or deterministic) connecting the reasons with the actions (a point I will expand on a bit later); and so there can be responsible actions for reasons which are not indeterministically caused by those reasons.

Why we should reject the indeterministic-causation account is, however, not a topic I propose to go further into here. For the rest of this paper I will just assume that that view is out of the running and focus on the issue between my view, simple indeterminism, and O'Connor's agent-causation view. I will look at and respond to O'Connor's objection to my view. And then I will give my objections to his view. In the end I will grudgingly allow that we may speak of agent-causation if we like but we should regard such talk as compatible with my view of what is really there.

⁷ See O'Connor, "Indeterminism and Agency: Three Recent Views," and O'Connor, "Agent Causation," in T. O'Connor (ed.), *Agents, Causes, and Events* (Oxford: Oxford University Press, 1995), pp. 171–200.

⁸ R. Nozick, *Philosophical Explanations* (Cambridge: Harvard University Press, 1981).

⁹ R. Kane, *Free Will and Values* (Buffalo: State University of New York Press, 1985).

But first let me state my view more fully. I will just state it without trying to argue for it, except for those points of contention between O'Connor and me.

Every action, according to me, either is or begins with a causally simple mental action, that is, a mental event that does not consist of one mental event causing others. A simple mental event is an action if and only if it has a certain intrinsic phenomenal quality, which I've dubbed the "actish" quality and tried to describe by using agent-causation talk radically qualified by "as if": the simple mental event of my volition to exert force with a part of my body phenomenally seems to me to be intrinsically an event that does not just happen to me, that does not occur unbidden, but it is, rather, as if I make it occur, as if I determine that it will happen just when and as it does (likewise for simple mental acts that are not volitions, such as my mentally saying "Shucks!"). A simple mental event's having this intrinsic actish phenomenal quality is sufficient for its being an *action*. But its having the quality entails nothing either way as to whether it satisfies the incompatibilist requirement for *free* action (which is that it not be causally necessitated by antecedent events).

An action may be causally complex, may consist of a simple mental action plus consequences of it. For example, my action of voluntarily pushing with my arm and hand against a door begins with a volition, a simple mental act of willing to exert a certain force in a certain direction with my arm and hand, and consists further in that volition's causing my arm and hand to exert such a force. My action of opening the door has a still further component of the door's opening being caused by the force exerted against it by my arm and hand.

Now, as I explained earlier, if an event is *not* an action of mine – for example, the door's opening – then I can make that event occur only by causing it, that is, by performing some action that causes it. But I make my own free, simple mental acts occur, not by causing them, but simply by being their subject, by their being my acts. They are *ipso facto* determined or controlled by me, provided they are free, that is, not determined by something else, not causally necessitated by antecedent states and events.

O'Connor readily concedes that there is an actish phenomenal quality to any simple mental action – a quality of its seeming as if I directly bring it about. His difficulty with my account of an agent's determining her own action is that he cannot see how the mere presence or absence of the actish phenomenal quality, without there being also a genuine causal relation to the agent, can make the difference between an event's being one that the agent determines and its being one that the agent does not determine. He says:

... how can it plausibly be maintained that the fundamental actions in virtue of which we have control over events within and immediately external to our body may themselves be simply uncaused occurrences; that we may be said to be in control of *these* simply by virtue of their obtaining (undeterminedly)? Perhaps the challenge is most directly posed in these terms: why are volitions *intrinsically* such as to confer control in the absence of determining antecedents?¹⁰

The issue, O'Connor says, is

whether or not control over one's actions obtains in virtue of certain purely intrinsic, non-relational features of those actions together with the negative condition of there being no determining prior causes. I am inclined to think that the onus is clearly on one who affirms this thesis to explain why we should think this to be so.¹¹

I confess that I can manage to feel some sympathy for O'Connor's uneasiness here. I am not completely baffled by it. I understand the temptation to say that the actish phenomenal quality alone cannot be enough, the temptation to take literally the impression that the actish quality consists in (the impression that I directly cause my mental action) and treat this impression as an awareness of a real causal relation between me and my action. I appreciate the temptation and I would be more prone than I am to succumb to it were it not that, first, the positing of this special sort of causal relation does not seem to me to solve the problem alleged for my view and, second, there seem to me to be unacceptable difficulties in the very concept of a direct causal relation between the agent as such and a mental event of hers.

The alleged problem with my account could, it seems, be put this way: we feel that an agent's determining or controlling her action – making it occur – must involve something more than just the undetermined occurrence of an event with certain intrinsic properties, whatever those intrinsic properties might be. If that is a fair way of putting the difficulty, then O'Connor's own agent-causation account seems to fail to avoid it.¹² He holds that where an agent causes a simple mental event *e* it is not this event *e* that is her action, but rather it is the whole causally complex event of *her causing e*. He says that "rather than there being a causal relation between agent and action, the relational complex constitutes the action."¹³ He tells us that the event the agent directly causes in an action is the coming to be of a state of intention.¹⁴

On this account it seems that an agent controls her *action* simply in virtue of the action's having a certain intrinsic property (its being the

¹⁰ O'Connor, "Indeterminism and Free Agency: Three Recent Views", p. 505.

¹¹ *Ibid.*, pp. 505–506.

¹² O'Connor, "Agent Causation."

¹³ *Ibid.*, p. 182.

¹⁴ *Ibid.*, p. 198, note 15.

agent's causing something) and its being causally undetermined. About this account one might ask, echoing the question O'Connor asks about my account: Why should an action's being undetermined and having a certain intrinsic property be enough to make it an event that the agent determines or controls? If her action is not caused by anything, not even the agent, then how can we hold the agent responsible for it?

O'Connor recognizes this sort of objection.¹⁵ He tries to rebut it by first arguing that "the very idea of there being sufficient causal conditions for an agent-causal event is unintelligible"¹⁶ and then saying of an agent-causal event:

Its very nature precludes the possibility of there being a sufficient causal condition for it (as I argued earlier), being an event that is the agent's causing the event internal to it (*e*). Now the event *e* is itself clearly under the control of the agent, since *he* caused it (directly). But would it not, then, be perfectly absurd to raise a doubt concerning whether the agent controlled *his causing e*? Indeed, it seems to me that the question of whether the agent has control over this event is ill framed – it is simply an instance of an agent's exercising direct control over another event.¹⁷

Here I find myself puzzled. I don't see why it would be absurd to suppose that we are entitled to say that in causing *e* the agent exercised control over *e* only if we are also entitled to say that the agent exercised control over her causing *e*. And I don't see why the question "How is it that the agent controls her causing *e*, given that her causing *e* was not causally determined by anything?" is any more (or less) absurd than the question "How is it that the agent controls her being the subject of an event with the actish phenomenal quality, given that this event was not causally determined by anything?" If the latter question is embarrassing for my account – a question that the account should answer but is unable to answer (as O'Connor thinks it is) – then the former question should be embarrassing for O'Connor's account; and contrapositively, if the former is *not* embarrassing for O'Connor's account – not one that it needs to answer – then neither is the latter embarrassing for my account (needless to say, it is the components of this last conditional that I am inclined to affirm: neither question is one that the account it pertains to needs to answer).

That O'Connor thinks otherwise must be because he holds the following about agent-control and causation: (a) if an event is such that it is possible for it to be caused, then an agent's controlling or determining it must be a matter of its being caused in an appropriate way, but (b) if an event is

¹⁵ He quotes Chisholm's remark, "If we say this [that the agent's causing *e* was not caused by anything], then we cannot hold *him* responsible for his causing *e* to happen." (Chisholm, "Reflections on Human Agency," *Idealistic Studies* 1 (1970), p. 40.)

¹⁶ O'Connor, "Agent Causation," p. 185.

¹⁷ *Ibid.*, p. 187.

such that it is *not* possible for it to be caused then, of course, an agent's controlling or determining it cannot be a matter of how it is caused. (b) is obvious enough. But (a) is not obvious, and O'Connor has offered no argument for it. My intuition that it is obvious that an event's having the actish phenomenal quality and being causally undetermined makes it one that is controlled and determined by the agent seems to me just as plausible as O'Connor's intuition that it's obvious that an event's being an agent-causal event and therefore one that *could not* be causally determined makes it one that is controlled by the agent.

So my first point is that I don't see why an event's being an agent-causal event and necessarily lacking a sufficient cause makes a better sufficient condition for the event's being one the agent controls than does the event's having the actish phenomenal quality and contingently lacking a sufficient cause. And, I might add, my proposal as to what grounds our saying that an agent directly determines an event has the advantage over O'Connor's of not requiring us to add to our basic posits a special *recherche* sort of causation.

But I also do not find compelling O'Connor's argument that an agent-causal event must necessarily lack a cause, that it is unintelligible to suppose that an agent-causal event might be causally determined by its antecedents (granting for the sake of this argument that the notion of an agent-causal event is itself intelligible). O'Connor's argument for this claim is the following:

In general, that which is causally produced in the first instance is always an event or state having a causally *simple* structure: an object *O*'s exemplifying *intrinsic* properties p_1, p_2, \dots at time t_0 . Causally complex events can also be caused, of course, but only in a derivative way: where they have the form *event X's causing Y*, whatever causes event *X* is a cause *thereby* of *X's causing Y*. In the special case of an *agent's* causing an event internal to his action, however, there is no causally simple component event forming its initial segment, such that one might cause the complex event (*S's causing e*) in virtue of causing *it*. Therefore, it is problematic to suppose that there could *be* sufficient causal conditions for an agent-causal event.¹⁸

This is an ingenious argument. But its premise is open to question. The premise is that what is causally produced in the first instance is always an event having a causally simple structure. I agree with the weaker proposition entailed by this premise, that what has a sufficient cause, in a direct and non-derivative sense, is always an *event*, never an enduring entity: if we ever speak of the cause of an enduring entity ("The potter's activity was the cause of the pot") we mean the cause of the event of that entity's coming to be. But the further implication of O'Connor's premise, that only causally *simple* events can be caused in the most direct sense, is not evident

¹⁸ Ibid., p. 186.

to me. This may be a plausible claim for the case of event-causation, where the cause-end of the complex event is itself an event. It may well be that, where X is an event, the only basis there can be for saying that something was a sufficient cause of X 's causing Y is that it was a sufficient cause of X and X caused Y . But why must a similar thing hold for the very different case of agent-causation, for the case of X 's causing Y where X is an agent and not an event? It's not clear to me why the principle that causally complex events cannot be directly caused must hold for agent-causation events if it holds for event-causation events. It needs further argument that there cannot in the special case of agent-causation be a direct cause of the complex event, and O'Connor has not supplied any.

It is clear, by the way, why the agent-causation account does not want to allow that any extrinsic condition is needed to give the agent control over her causing e . For if it did allow this, then it would be in serious trouble. The only extrinsic condition the agent-causation account has to suggest – the only thing consistent with its idea as to what agent control consists in – is that we posit a further agent-causal relation, this time between the agent and the event of her causing e . But the same question can again be raised about the agent's causing of her causing e : what gives the agent control of *that*? Thus the account would either fail to give us any ultimate answer to the question of what agent-control consists in or else give us the very implausible answer that it consists in an infinite regress of agent-causings.

Those are my reasons for thinking that positing agent-causation does not solve the alleged problem it is meant to solve. In answer to the question "Why should an action's being undetermined and having a certain intrinsic property be enough to make it an event that the agent determines?" the agent-causationist, as much as the simple indeterminist, has to say, "Well, it just is"; and the property of agent-causation gives no better justification for saying this than does the actish phenomenal quality. In fact, it may give a worse justification, for there are reasons – two I can think of – to worry about the intelligibility of the notion of agent-causation.

One of my worries on this score is inspired by a question C.D. Broad raised about the coherence of the agent-causation theory. Broad asked:

How can 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?¹⁹

On the agent-causation 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

¹⁹ C.D. Broad, *Ethics and the History of Philosophy* (London: Routledge & Kegan Paul, 1952), p. 215.

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?

O'Connor confronts this worry and responds by saying:

An agent-cause does not produce a certain effect by virtue of its very nature, as does an event-cause, but does so at will in the light of considerations accessible to the agent at that time. And so a full explanation of why an agent-caused event occurred will include, among other things, an account of the reasons upon which the agent acted.²⁰

O'Connor seems to be suggesting here that when we have an agent-caused event the explanation of the specific nature and timing of the event – that which makes it right to say that the agent determined that just that sort of event rather than some other sort would happen at just that time rather than at some other time – lies in the fact that the agent caused the event for certain reasons she had for bringing about that sort of event at that time.

For this suggestion I see a couple of problems. First, what about cases of spontaneous action where an agent acts in a certain way for no particular reason, at any rate not in order to satisfy any antecedently existing motive? Surely, agent-causation must be involved in such spontaneous actions if it is involved in any.

Second, and more important, it is a fact about reasons explanations²¹ that a full explanation in terms of the agent's reasons for acting as she did need not explain why there occurred that particular action *rather than* some other for which she also had reasons. An obvious case of this is where an agent is indifferent between alternative means to the same end – has no reason to prefer one over another – and just arbitrarily chooses one. For example, my reason for picking up the telephone was that I wanted to make a call. But that reason does not explain why I used my left *rather than* my right hand to pick it up, and indeed I need not have had any reason for using one hand rather than the other. Similarly, my reason for picking up the phone does not explain why I picked it up precisely when I did rather than a few seconds earlier or later, and I need not have had any reason for choosing that precise time rather than a slightly earlier or later one. It is possible that there was nothing at all that explains why the

²⁰ O'Connor, "Agent Causation," p. 184.

²¹ See Reid, *Essays on the Active Powers of the Human Mind*, pp. 285–286.

one thing was the case rather than any alternative (a possibility that I think O'Connor would not deny). But in that case, it seems natural to infer, there was nothing that *caused* the one rather than any alternative.

But it appears that the agent-causationist cannot make this natural inference. For surely, if agent-causation is involved in the action, it must be that *I caused* it to be my left rather than my right hand that I willed to move and *I caused* the action to happen precisely when it did rather than at a slightly different time. Thus the agent-causation theory is committed to supposing that a *cause* of something can fail to provide any *explanation* of it. If I am the cause of its being my left rather than my right hand that I will to move, but I am not the cause in virtue of any property or change in me at the time – I could in the same circumstances equally well have caused it to be my right hand for the same reasons I caused it to be my left – then there is nothing about this cause of its being my left rather than my right hand that I willed to move – me – that *explains* why it was my left rather than my right hand that I willed to move. If I am the cause of the precise timing of my volition, but it is not the case that I am the cause of this in virtue of any property or change in me at that precise time – I could in the same circumstances equally well have caused it to occur a bit later for the same reasons I caused it when I did – then there is nothing about this cause of its being at that precise time that *explains* its being at that precise time. 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. I don't know that I want to go so far as to say that this is self-evident, but its denial is sufficiently puzzling that we ought not to accept it without having compelling reason to do so.²²

²² What about indeterministic causation? Might it be causation without explanation? I think not. Let us look at an example. Suppose we have set up the famous two-slit experiment with photon detectors at each slit. A photon detected at slit *A* causes a green light to go on, and one detected at slit *B* causes a red light to go on. A photon is fired toward the screen containing the slits. This event, let us suppose, causes the red light to go on; but it does so indeterministically, since, according to quantum theory, it is compatible with the laws of nature that exactly the same event in exactly the same circumstances should have caused the green light to go on. So, one might be tempted to conclude, the photon's firing caused the red light to go on but does not explain its going on. But this would be wrong. What the firing of the photon caused, but also explains, is (a) the red light's going on *rather than neither light's going on*. What it does not explain is (b) the red light's going on *rather than the green light's going on*. But neither does it cause this. If one asks, "What caused it to be the case that the red light rather than the green light went on?" it would not be correct to answer, "the firing of the photon." Indeed, the correct answer seems to be that *nothing* is causally responsible for its being the red rather than the green light that goes on

My other worry about the concept of agent-causation is that agent-causation appears to be undetectable. Given that we have a person, *S*, who is the subject of a simple mental event that has the actish phenomenal quality – a volition, say, or mentally saying a word – and this event is not causally determined by antecedent events and states, what detectable difference would it make to the situation whether the agent-causal relation between *S* and that event was present or not? What independently specifiable difference could distinguish between the case where the agent causal relation is present and the case where it is not? That is, what independently specifiable feature could constitute the agent-causal relation? O'Connor does not try to give any account of this and it is hard to see what account could be given. However the person and her actish mental event are realized in the ultimate constituents of reality, it is difficult to imagine what additional thing might be realized there that would *force* us to describe it as a causal relation between the person as enduring substance and the event. Whatever independently specifiable difference might be pointed to, we would be free to regard its coming about as *another event* of which the person is the subject rather than as a brute causal relation between the subject and the actish event.

O'Connor does acknowledge that there might be concern about the detectability of the agent-causal relation.

... it seems that it is impossible in principle, for us ever to know whether any events are produced in the manner that the agency theory postulates, because such an event would be indistinguishable from one which was essentially random, not connected by even probabilistic laws to events preceding it.²³

His response to this worry is not to argue that agent-causation is detectable, but rather to claim that, however undetectable it might be, we are nevertheless forced to posit it in order to give an adequate account of reasons explanations of action (given that we reject an event-causation account invoking probabilistic laws). He thinks that my account of reasons explanations – which invokes neither agent-causation nor probabilistic laws – won't do.

My account of reasons explanations has it that when an action is correctly explained as one the agent did for a certain reason supplied by an antecedent motive (a desire or intention), the explanatory connection between the motive and the action is forged, not by causal laws (probabilistic or deterministic) or by agent-causation, but by an *intention* concurrent

(because nothing is causally responsible for its being slit *B* rather than slit *A* that the photon goes through). So we do not have here any state of affairs for which there is a cause of its obtaining that fails to explain its obtaining.

²³ O'Connor, "Agent Causation," p. 195.

with the action. The intention has the following content: it refers directly to the current action and to the remembered prior motive and says that this action is to satisfy that motive (or to help to do so). For example, if I push against a door in order to satisfy my antecedent desire to know whether it is locked – my reason for pushing against it is that I desire that information – then it is sufficient for the truth of this explanation that concurrently with my pushing I remember that desire and intend that this pushing will contribute to satisfying it.

But O'Connor thinks that this is not sufficient.

... reasons explanations require a mechanism of control that 'hooks up', so to speak, the agent's reasons and consequent decision (and action) On the agency [i.e., agent-causation] theory, an agent's capacity directly to produce a decision in the light of consciously held reasons fills the bill. We cannot *simply* appeal, as, for example, Ginet does, to *internal* (and referential) relations between concurrent intention and prior motives, on the one hand, and that same concurrent intention and the decision (or action), on the other. Without the mediation of a (necessarily causal) 'mechanism of control', prior motives cannot *explain* a decision, even though (as it happens) they may coincide with it.²⁴

I protest that I simply fail to see why the relation between prior motive and action that I specify does not suffice to guarantee the explanatory connection that we imply when we say the agent did the action in order to satisfy that motive. Why does O'Connor think that we need something further of a causal nature? (The term "causal" could be extended to any sort of explanation, I suppose, on the ground that it is always appropriate to state an explanation using the word "because"; but O'Connor is using the term in a narrower sense).

O'Connor suggests that something causal is needed in order to distinguish "S did A in order to satisfy her desire" from "S did A knowing that she would thereby satisfy her desire."²⁵ I agree that the latter does not entail the former, but my sufficient condition captures the difference. Where it is the case that S does A believing that she will thereby satisfy her desire but not the case that she does A in order to satisfy her desire, my sufficient condition does not obtain: though S *believes* of her action that it will satisfy her desire she does not *intend* of it that it will do so (though it is doubtful that one can intend of one's action that it will satisfy one's desire without believing of it at least that it might do so, it is certain that one can believe of one's action that it might satisfy one's desire without intending of it that it do so).

If O'Connor were still to insist that explanation always requires causation, then he would seem to me to be just clinging to a dogma, blind to the

²⁴ Ibid., p. 195.

²⁵ Ibid., p. 193.

possibility that in reasons explanation we have a fundamentally different kind of explanation, a non-causal kind.

Another reason O'Connor has for wanting to posit agent-causation (a reason we all have for being tempted to posit it) is that doing so would allow us to take the actish phenomenal quality seriously. It would allow us to take this its-seeming-as-if-I-directly-cause-it as a literal perception of reality; it is this that makes so prereflectively appealing the idea that what makes a person the agent of an actish event is that she directly causes it. So O'Connor might want to say that, if agent-causation is not independently detectable, then it must *supervene* on the facts that an event with the actish quality occurs in an appropriately complex creature and is not causally determined by its antecedents (and he should add, I think, that it supervenes *only* on such facts: I can't see what alternatives could be plausibly deemed sufficient for the presence of agent-causation). But one who says this has, it seems to me, come over to my view of what is really going on in free and responsible action, retaining only the *language* of agent-causation to describe it. One has, in effect, explicated being agent-caused as having actish quality plus being undetermined by antecedents.

This is agent-causation that I understand better and find more congenial. I might still object that to say that the agent causes the actish event is to stretch the meaning of the term "cause" a bit far, in that this sort of cause would not be any sort of explanation. If the champion of agent-causation replies that this is a small price to pay for the benefits of taking agent-causation talk seriously, then I'm inclined to say: OK, if you insist. I'll give you the agent-causation label as long as you give me my account of the reality it labels.

Department of Philosophy
Cornell University
Ithaca, NY 14853-3201
USA