

Why Agent Causation?

Timothy O'Connor
Indiana University

I. INTRODUCTION

The question of this paper is what it would be to act with freedom of the will. What kind of control have we in view (albeit inchoately) when we speak, pretheoretically, of being “self-determining” beings, of “freely making choices” for consciously considered reasons (pro and con)—of its being “up to us” how we shall act? My question here is not whether we have (or have any reason to think that we have) such freedom nor what the most robust account of our freedom compatible with late-twentieth-century science is. Many contemporary philosophers are all too ready to settle for a deflationary account of freedom and declare victory, with some brief remarks reminding us that we were created a little *lower* than the angels. I am not so sanguine about the ability of such accounts to leave reasonably intact our judgments about human autonomy, dignity, and responsibility. But, as I’ve said, that’s not my concern here. Instead, I want to revisit the question of what exactly “self-determination,” on our ordinary conception of it, comes to.

I will assume, for familiar reasons, that freedom of will requires that our actions be causally undetermined.¹ Recently, some philosophers have advanced sophisticated accounts that incorporate this condition of significant

causal indeterminism, some of which are refinements of familiar views and some of which are novel.² Such views fall into three basic camps: agent-causation accounts, simple indeterminism, and causal indeterminism. (Each of these positions will be explained below.) In an earlier article, I argued that simple indeterminism and causal indeterminism are inadequate approaches—doing so through an examination of the three best contemporary defenses of them—and concluded that incompatibilists must revive the traditional notion of agent causation, despite the well-known objections that have been raised to it.³ Subsequently, I have begun the task of refurbishing and defending this idea.⁴

But proponents of the alternative views have not been without their replies. (Alas, such unseemly displays of willful obstinacy are common in philosophy.) So in the present paper, I want to explain as carefully as I can why the incompatibilist who does not take free will to be an outright incoherent idea⁵ should believe in agent causation—unfashionable though this may be.

II. AGENT CAUSATION BRIEFLY EXPLAINED

To be an agent causationist, you have to believe in a nonreductionist account of causation generally. “Ordinary” causation by events or states of affairs may conform to sharp, generalizable patterns—and it may even fall out of one’s account of such causation that there will be such patterns—but causation, for the nonreductionist, is not constituted by such patterns. The core notion of causal production (or in a more technical jargon, “causal oomph”) is the notion of an ontologically basic sort of relation.

But that a reductionist analysis of causation is not possible does not rule out all attempts at explication. Here, in broad strokes, is a nonreductionist explication of *event* causation of the sort that I favor. Ordinary objects possess a range of causal powers. When placed in the appropriate circumstances (circumstances that either simply remove impediments to the action of an object already in a state of readiness to act or stimulate a latent causal mechanism), they will exercise these causal capacities. These capacities have their basis in the object’s (or system of objects’) underlying nature—its structural and dynamical properties. Such properties are thus “two-sided”—they have both a categorical (e.g., being sharp-edged) and a dispositional (e.g., bestowing a capacity to cut certain kinds of object) nature.

How should the notion of agent causation be understood by one who accepts this basic account of event causation? Let us first note that the notion of a particular agent freely and directly (i.e., agent-causally) bringing about an effect requires that the particular is an agent capable of representing pos-

sible courses of action to himself and having certain desires and beliefs concerning those alternatives.⁶ What sort of events do agent-causes bring about? As I will think of it, these agent-caused internal events are determinate (immediately executive) intentions to act in various ways. The agent-causal relation itself is conceived simply as follows: Wherever the agent-causal relation obtains, the agent bears a *property* or set of properties that is “choice-enabling” (i.e., in virtue of such properties, the agent has a type of causal power which, following Reid, we may term “active power”). But this “active power”—the causal power in virtue of which one has freedom of will—is not characterized by any function from circumstances to effects (as is the case with event causal powers). For the properties that confer such a capacity do not themselves (in the appropriate circumstances) necessitate or make probable a certain effect. Rather, they (in conjunction with appropriate circumstances) *make possible* the direct, purposive bringing about of an effect *by the agent* who bears them. Such properties thus play a different functional role in the associated causal process. Though the very same relation of causation is involved, these properties give rise to a fundamentally different type of causal power—one that in suitable circumstances is exercised at will by the agent, rather than of necessity, as with objects that are not partly self-determining agents.

There is more to be said and asked about such an account of freedom of will (and we will ask some of those questions below), but it will have to suffice for now as a basic picture. It is clear enough why one might be initially attracted to it. For, on such an account, rather than my activity’s being a product of external conditions that impinge on me in various ways, establishing internal states that in turn cause the behavior, I am quite literally the cause (source, point of origination) of my own behavior.

But appearances can be misleading. Some incompatibilists have recently urged that such an appeal to a primitive form of causation is either useless or positively incompatible with a basic feature of action that we have been silent about up to this point—viz., that actions are explainable by reasons the agent had in acting as he did. So let us now consider such arguments and the alternative accounts of freedom of will they are said to bolster.

III. SIMPLE INDETERMINISM AND AGENT CAUSATION

The most prominent contemporary defender of simple indeterminism is Carl Ginet. So let us use his account as representative of the general view. The central feature of Ginet’s theory of action is the claim that there is a *simple* mental action (lacking internal causal structure) at the core of every causally complex action.⁷ In some cases, such simple mental actions are complete in

themselves, not being part of a more complex action, as in mentally saying a word or forming an image. Simple mental acts of this sort differ *intrinsically* from passive mental occurrences such as unbidden thoughts or images in virtue of what Ginet terms an “actish phenomenal quality”—the quality one might characterize (in the case of mentally saying a word) as “its seeming as if I directly produce the sound in my ‘mind’s ear.” Ginet emphasizes that this quality “belongs to the manner in which the word occurs in my mind and is not a distinct phenomenon that precedes or accompanies the occurrence of the word.”⁸

In the central sort of case, however, this simple mental action will be a volition that constitutes the core element of one’s voluntarily exerting the body, causally producing that exertion.⁹ The intentional content of the volition is that you are voluntarily exerting your body in a specific sort of way. This content, though, is directed to the immediate present (thereby differing from decision or intention), making it a “fluid” sort of mental activity over time.¹⁰

As a simple indeterminist, Ginet holds that such bodily actions are free if the simple mental actions at their core are uncaused. But this seems unsatisfactory to many (agent causationists included) for the following simple reason. If a simple mental action of mine is uncaused, if it is in no sense determined to occur by anything at all, then it is not determined to occur by *me* in particular. But if I do not determine its occurrence, then it is not under my control.

Ginet’s response to the agent causationist who levels this charge is as follows. Consider the *agent causationist’s* account of a basic action. It is an uncaused event of the form *agent S causes event e*. But if the lack of causation of the simple indeterminist’s volition renders it uncontrolled by anything, including the agent himself, then the same must be true in the case of an uncaused agent causing.¹¹ So either both accounts fail to adequately characterize self-determination or the idea that an agent’s control over her own activity must involve causation is mistaken. But if the latter is the case, there is no reason to complicate our picture of agency with the notion of agent causation.

However, this *tu quoque* reply masks an essential difference between the two sorts of events being posited. Both sorts are held to occur uncaused. But there is internal causal structure to the agent-causal event that is lacking in a simple volition. And this difference in causal structure bears directly on the question of whether their respective agents control their occurrences. An agent-causal event is intrinsically a doing, an *exercise* of control. Ginet claims that this is true of uncaused volitions as well, in view of their “actish phenomenal quality.” However, “control,” “determination,” and allied notions cannot be grounded in intrinsic, phenomenal characteristics alone—they clearly require some sort of causal elucidation.¹²

One who does not favor Ginet's simple indeterminist view might pick up the argument at this point and insist that, however one wants to describe the agent-causal event itself ("an exercise of active power"), the agent causationist's *own* original line of reasoning suggests that one controls this event only to the extent that one causes it. So if it occurs uncaused, its occurrence is not something that was up to the agent, something he controlled.

I suggest that there is something suspicious about this demand for some further event by means of which one controls what is intrinsically an exercise of control. The suspicion that such a demand is unreasonable is confirmed, I think, by the following simple argument that shows that an agent-causal event *couldn't* be causally determined. For instantiations of causal *relations* (causally complex events) are not themselves directly on the receiving end of other causal relations—instead, instantiations of intrinsic properties (causally simple *states* or events) are. Causing is the *producing* of events, rather than what are *produced* (in the first instance). Compare an ordinary case of an event-causal process (consisting of event *F*'s causing event *G*) being caused by some further *event E*. Surely this can consist only in *E*'s causing *F*, the front-end relatum of the complex event.¹³ When I reflect on the matter, I cannot but regard this as reflecting an evident, *general* truth about causation. If this is right, then an *agent-causal* event could not be caused for the simple reason that the cause in this case is not itself an event.

Ginet, however, argues that while the above may be a plausible claim for the way in which an *event-causal* event is caused, it is not obvious for the case of an *agent-causal* event. But why would one hold this in the event-causal case, if not because it reflects a general conceptual truth about causation and relations? Let us set that question aside, though, in order to examine further the implications of this suggestion. Notice that what is being countenanced is that while events can never cause event-causal events, they might cause agent-causal events. It is, thus, a claim about *event* causation—about what kinds of entities events can bring about. And the claim is that events can directly bring about causal relations when they relate an agent to an event but not when they relate an event to a further event. If true, this is a very puzzling fact. What is it about the agent-causal relation, in Ginet's understanding of it,¹⁴ that explains this fact? Let "*A*" and "*B*" be events and "*S*" an agent, and consider the allegedly possible event *A*'s causing *S*'s causing *B*. Should we also allow that there could be a direct cause of the relation between *A* and *S*'s causing *B*, one that does not do so by causing *A* itself? Granted, this is now a case of an event's causing another event, but the second event is an instance of *agent* causation, so perhaps that makes it sufficiently special as to allow for the direct causation of a causal relation.

If my strategy has been successful, my reader will by now be inclined to agree that it is simply wrongheaded, in general, to countenance direct

causings of causal relations. Whether our causes be events or agents, *what* they cause are causally simple states. To suggest otherwise, for either sort of causation, is to render the original cause strangely inert or at least deficient. (Something else must bring about their *causing* of the event.)

So to reiterate my earlier conclusion, since agents are not causally simple states, agent-causal events cannot be caused. This conceptual truth lends credence to my contention that it is simply inappropriate, or confused, to suggest that an agent-causal event—a direct *exercising* of control by an agent—must itself be caused by the agent (or some state of the agent) if he is to exercise control over *it*.

Let us now consider a second objection to agent causation that is made by simple indeterminists. We explain intentional actions in terms of reasons the individual has for so acting. According to several philosophers, this simple truism poses a difficulty for the theory of agent causation.¹⁵

In order to understand the different forms this charge takes, we must first set out a fragment of an account of reasons explanation of actions from the perspective of the theory of agent causation. Consider, then, the explanation of action by reference to a prior *desire*. An agent causationist might give the following general conditions for the truth of such explanations:

- S* acted then in order to satisfy her antecedent desire that ϕ if:
- (i) prior to the action, *S* had a desire that ϕ and believed that by so acting she would satisfy (or contribute to satisfying) that desire;
 - (ii) her action was initiated (in part) by her own self-determining causal activity, the event component of which is the-coming-to-be-of-an-action-triggering-intention-to-A-here-and-now;
 - (iii) concurrent with the action, *S* continued to desire that ϕ and intended of the action that it satisfy (or contribute to satisfying) that desire;
 - (iv) *S*'s concurrent intention was a direct causal consequence of the action-triggering intention brought about by the agent, and it causally sustained the completion of her action.

Simple indeterminists have claimed that this explanation schema is inadequate in at least two different ways. Some claim that the second condition is perfectly idle, since the other conditions are perfectly sufficient to explain the action (thus vindicating the simple indeterminist's wholly noncausal account of reasons explanations).¹⁶ And Carl Ginet has recently made the stronger claim that agent causation is inconsistent with the fact that reasons explanations of actions are often incomplete in certain respects. I may have a reason to reach out and grab an airborne ball headed my way but no reason for using my right hand instead of my left (even though I do in fact use

my right hand). Similarly, I may have a reason for calling a friend on the phone but no reason for calling at the precise moment I do. Let us focus on this second example. The agency theory maintains that, if I was free in making the phone call, then I *caused* the state of intention that issued in the bodily movements. But because I was not *caused to so act* just then and *had no reasons* for acting just *then*, there was nothing that explained the timing of my action. And this, it is held, is problematic: how can a *cause* of an event's occurring at a certain time fail to *explain* its occurring at that time? This reveals, it is urged, that it is incoherent to appeal to the agent—simply qua agent, apart from particular features of him at the moment his causal power was exercised—as the cause of the state of intention's being formed.

I will consider these objections in turn. The first objection alleged that agent causation plays no real role in the explanation of the action. If one desired that ϕ and intended throughout the action that it satisfy that desire, then one has thereby acted for the purpose of satisfying that desire; no appeal to a special sort of action-starting event is necessary.

But if my earlier contention that the exercising of control in acting must have a causal structure is correct, then this objection cannot be right. It is not sufficient to explain an action that one have had prior desires or intentions whose contents “fit” the action performed; one must have controlled the performance of the action (for the very purpose of satisfying the desire or intention in question). The agent causationist need not hold that the *only* way in which exercising control for the purpose of satisfying a prior desire can be carried out involves agent causation. Agent causation is appealed to in order to explain freedom of action, not action simpliciter. (“Controlling” the course of one's activity is a weaker notion than acting with free will.) So while it is true, I think, that the agent-causal condition (*ii*) is not itself a necessary condition for acting in order to satisfy some desire, it is necessary that there be some causal condition or other that accounts for the agent's exercising control in acting as he does. Consequently, the agent-causal condition is not idle, and simple indeterminism has not been vindicated.

Let me now turn to Ginet's argument (inspired by some well-known remarks of C. D. Broad's) that the agent causationist is committed to the possibility of a cause of an event that does not explain the event. Here we have to be careful, I suggest, to identify just what event it is that is in question. Suppose, as the agency theory would have it, that I causally generate a state of intention to call my friend here and now. Let us call the event of that state of intention's coming to be (at time t) “ e .” I am (at t) the cause of e . Yet I, qua its cause, also figure into the explanation of e 's occurrence, insofar as I had reasons for so acting.

“But what about the event e 's occurring at t rather than at t_1 ? Ex hypothesi, there is no explanation of this event, and yet you caused it, according to

the agency theory.” Here I think Ginet is mistaken to hold that there is any such event as *e*’s occurring at *t* rather than at t_1 in addition to *e* itself. There is no reason to hold that corresponding to every contrastive *fact* about a contingent occurrence there is a distinct “contrastive event,” where an event is understood to be a concrete entity. We might recast the objection in this way: On the agent causation view, there is a feature of, or fact about, *e*—viz., that it occurred at *t* rather than at t_1 —that obtains in virtue of some causal factor that nonetheless does not explain that feature’s obtaining. But, so stated, the objection would appear to be an implicit rejection of the more general idea of nondeterministic causation. For consider a case of indeterministic event causation from quantum mechanics.¹⁷ A photon is fired at a screen with two slits. If detected at slit *A*, it will cause a green light to go on, and if detected at *B*, it will cause a red light to go on. Via an indeterministic process, it causes the red light to go on. Now it seems that in virtue of *the firing of the photon*’s causing the event *the red light’s going on*, the fact of *the red light’s going on rather than the green light’s going on* obtained. But we do not seem to have an explanation of this fact. If this scenario is incoherent, it is not obviously so. Likewise, there is no obvious incoherence in the agent causationist view that an agent causes a certain event (whose occurrence is explained by certain reasons he had at the time) without there being an explanation for every contrastive fact that obtains in virtue of the event’s occurring.

A final objection to agent causation from the simple indeterminists is epistemological, rather than metaphysical, in nature. It is that agent causation is undetectable, and so we could never have any reason to posit it. Evidence of significant indeterminism at the level of human action would not favor the hypothesis of agent causation over that of causal randomness.

In reply, I would begin by observing that whatever force this objection has, it is had equally by a parallel objection to a nonreductionist account of *event* causation. On such an account, causal relations are among the basic furniture of the universe. And like Ginet on agent causation, a Humean reductionist about event causation will say that all we have direct evidence of is the *pattern* of relations among types of events. Having noted this parallel objection, there are two things to be said in response to *both* of them. The first is that we don’t simply observe an apparently undetermined event (in the case of an agent’s activity) or a pair of events falling under a certain type of lawlike pattern (in the general case of event causation). Rather, we seem to directly observe their causal connectedness—at least in some cases. For example, I don’t merely observe the movement of the hammer and then the movement of the nail—I perceive the hammer’s *moving* the nail. And in the (putatively agent-causal) case of my own deliberate formation of intentions, the event doesn’t seem merely to occur but seems to be my directly exerting causal control in bringing it about. (Note Ginet’s own frank admission of the

“actish phenomenal quality.”) Such apparent perception of causality could be mistaken, of course, but by the same token, so could our apparent perception of ordinary physical objects, and few would take that as a reason for disbelief.

The second thing to be said about the objection is that there are theoretical, as well as perceptual, reasons for accepting the existence of event-causal relations in the actual case and agent-causal relations in a hypothetical case involving evidence of significant (and apparently non-law-governed) indeterminism in the production of human behavior. For the alternative to positing causal relations of the appropriate sort is to accept such occurrences as brute, unexplained facts. In my book, this is an unreasonable stance.

I have now completed my discussion of arguments favoring simple indeterminism to agent causation. Before I move on, though, I want to make one final point about the simple indeterminist view. The simple indeterminist holds that a free action must be causally undetermined and that there *need* not be any causation (of even an indeterministic variety) involved in the generation of such actions. But Ginet, for example, appears to allow that a given free action *might* be caused in an *indeterministic* fashion. My point is that the simple indeterminist cannot, in consistency, allow for this. For on this view (as on the agent-causal view), an agent’s determination of his own decision, his control of it, is an *intrinsic* feature of the decision itself. But if a decision were indeterministically caused by the agent’s reasons, say, then the agent’s determination of his action would reside in the causal activity of (the state of his having) those reasons and not in the intrinsic features of the decision itself. Hence, the simple indeterminist’s analysis of what it is to freely control the formation of one’s decision leads to the necessity of supposing that such decisions are entirely uncaused. Without pausing to develop an argument here, I will simply assert that, given the fact of (probably indeterministic) causation at the fundamental levels of matter, entirely uncaused decisions would seem inconsistent with a materialist (or even an emergent dualist) account of human mental activity, and so would seem to force an acceptance of substance dualism of a very strong sort.

IV. CAUSAL INDETERMINISM AND AGENT CAUSATION

Causal indeterminism is an alternative to both agent causation and simple indeterminism as an account of free will. According to the causal indeterminist, our free decisions are caused but nondeterministically. And the most salient of the causal conditions of those decisions are the states of having reasons for deciding in that manner. In many circumstances, persons have (at least partially) distinct desire-belief complexes (more simply, “reasons”)

that point toward different courses of action, and the performance of *any* of these—*not* just the one that was actually done—would “graft” coherently onto precisely the same prior circumstance. Each such reason is a potential cause of the corresponding behavior; whichever action is (as it happens) undertaken will have been caused by the matching reason or reasons.

The basic motivation for this view is that it provides a way of recognizing (contra the simple indeterminist) that control over one’s decision-making must be causal in nature, while avoiding commitment to an irreducible form of agent causation and its attendant problems, such as explaining how the two forms of causation could systematically interact within a single being. But this basic advantage comes at the cost of an equally basic problem, viz., explaining how it can be up to *me* on a given occasion that certain of my reason states exert their indeterministic propensity to cause a decision to act, while others do not (though they might have done so).

A decade ago, in his book *Free Will and Values*, Robert Kane produced a causal indeterminist account of free will intended to deal with this issue.¹⁸ And in a series of subsequent articles, he modified and refined the original account in response to various criticisms. Elsewhere, I have criticized the version of his account sketched in the most recent of these articles.¹⁹ However, in an impressive new book, Kane has extensively developed his basic account in several respects, which I will now explain in basic terms and then assess.²⁰

On Kane’s account, freedom of the will is *directly* manifested only in the context of conscious deliberation. Such deliberation normally is terminated by “choices” or “willings” that form intentions and create purposes that guide actions, now or in the future.²¹ Freedom of the will is constituted by undetermined “self-forming willings” by virtue of which the agent is the ultimate creator of some of his own purposes. Kane identifies at least six different kinds of such willings:²²

- 1) Moral choices
- 2) Prudential choices
- 3) Practical judgments and choices
- 4) Efforts of will sustaining purposes
- 5) Attentional efforts directed at self-control and self-modification
- 6) Changes of intention in action

(1)–(3) are the basic categories of choices about what is to be done, either now or in the future. The cases falling under the heading of (3) are simply those deliberative situations in which duty or desire and short-term or long-term interest-conflicts do not play a significant role. (4)–(6) are kinds of effort

that may be involved in the deliberative process leading to a decision to act. In what follows, I mainly focus on moral and prudential choices, which are treated as essentially alike on Kane's account, and on the efforts of will that precede them.

The basic structure of moral and prudential deliberations is held to be as follows.

First, the agent is in a divided motivational state, feeling inclined to pursue each of two or more incompatible courses of action. The different motivations for each such action are incommensurate, reflecting different ends or purposes.

Second, such conflicts result in an "effort of will"—a struggle by the agent to get his ends or purposes sorted out.²³ In a context of moral decision-making, they are struggles to act in accordance with perceived obligation in the face of contrary motives to act differently, with the outcome being uncertain in the agent's mind prior to the moment of choice. It is supposed that an effort of will is an *indeterminate* (and not merely undetermined) process, analogous to the indeterminacy of the position and momentum of an individual particle at the quantum level prior to a measurement. Kane suggests that this could be true of even complex processes in our brains involving many neuron firings and connections, provided that there be "chaotic processes"—processes sensitive to minute changes in initial conditions—in the brain that could amplify indeterminate events at the microlevel. Persons experience these complex processes as efforts of will they are making to resist temptation in moral and prudential situations.²⁴

Third, resolution of the indeterminacy is achieved through the agent's choice, which, because of the indeterminacy of the process leading up to it, is a causally undetermined event.

Finally, we should note that, according to Kane, it is

because their efforts are thus a response to inner conflicts embedded in the agents' prior character and motives that their characters and motives can explain the conflicts and why the efforts are being made, without also explaining the outcomes of the conflicts and the efforts. [They] provide reasons for going either way, but not decisive reasons explaining which way the agent will inevitably go.²⁵

This thumbnail sketch of Kane's basic picture is enough for us to begin raising questions about it. As I do so, I will introduce further wrinkles that Kane adds to address many of these questions.

On Kane's account, then, a moral or prudential choice is an undetermined, causal outcome of an indeterminate struggle or effort to sort out one's purposes via deliberation. One basic problem with this account, as I see it, concerns how this choice occurs. Consider the simpler quantum analogue. An isolated particle (e.g., an electron) moves toward a thin atomic barrier.

Its position and momentum are not both determinate; rather, it is in a superposition of states (i.e., instead of having a fully determinate position, its position is describable only by a particular *range* of values) such that there is a certain probability that it will penetrate the barrier. But now, on the orthodox interpretation, the resolution of the indeterminacy is a consequence of a quantum mechanical measurement event. This, of course, raises puzzling questions of its own. However, my purpose is not to raise those questions but instead to note the apparent need for a special kind of “trigger” event to induce *determinacy* in the system. It’s hard to see how Kane can accommodate this need within his picture of free choice. For if the choice which induces determinacy is disconnected from the indeterminate process itself, it will not admit of a causally indeterministic explanation in terms of that process’s antecedents. This will either land us back into simple indeterminism or push us to admitting a *sui generis* agent-causal activity, which causal indeterminism is supposed to enable us to avoid.

I note in passing that, for Kane, *practical* choices (ones not involving conflicts between duty or long-term interest and other sorts of motivation) do not appear to be indeterminacy-resolving events, at least in many cases. In practical deliberation, there is indeterminacy at the outset, connected with the intentional efforts to open oneself up to all relevant considerations and to continue deliberating. But he allows that often the outcome of such deliberation may be causally determined by its immediate antecedents, once it becomes clear what the best course of action is. (Might the event of coming to believe that option *X* is the best alternative be an analogue of indeterminacy-resolving moral choice?)

But, returning to moral and prudential choice, perhaps its nature can be explicated in a way that leaves intact the basic causal indeterminist picture. There is still the fundamental question of how it is that I may be in control of the way it goes in a given case. How is it up to me that, on this occasion, this one among two or more causally possible choices was made?

Kane thinks that part of an answer to this can be given by positing a “self-network.” This is conceived as a comprehensive network of neural connections representing the agent’s general motivational system—the plans, aspirations, and ideals in terms of which she defines herself as an agent and practical reasoner.²⁶ (These diverse neural circuits could act in concert by forming a synchronized pattern of oscillations or wave activity of a sort that has been suggested in recent studies.) He suggests that

the neural events corresponding to our efforts and choices would be overlaid by the wave patterns unifying the self-network—so that the wave patterns and the effort or choice events are coupled, causally influencing and interacting with each other. The effort and choice events would occur, so to speak, “within” the self-network whose distinctive wave patterns of oscillations were

superimposed on them. In turn, the superimposed patterns of oscillations of the self-network would be contributing causes to choice, pushing one competing reason-network over the top, so to speak, so that A is chosen for reasons R rather than B for reasons R' (or vice versa)—thus supporting the belief that the efforts and choices are our doings, the products of our selves.²⁷

So it is the influencing activity of the superimposed self-network that turns back the objection that I am unable to directly determine which outcome my reason states will give rise to in a given case. I do so determine this because this self-network essentially constitutes who I am, psychologically speaking. The self-network, then, is a kind of surrogate for traditional appeals to an irreducible causation by the agent, qua agent.

However, when we consider the role of the self-network, precisely the problem it was supposed to solve in relation to the indeterministic activity of my reason states—i.e., my apparent inability to directly determine the outcome of the chancy, indeterministic events constituting my reason states—reappears in relation to *it*. For its activity, too, is held to be indeterministic, potentially giving rise to any of a range of choices in a given context. Insofar as the outcome will be a causal consequence of this network acting on the states constituting my deliberation, I exercise a form of control over it. But the question is whether it is a form of control that is sufficient to ground personal responsibility. On the agent-causal view, I myself directly determine whatever outcome occurs. Unless Kane's contrasting view is coupled with a reductionist account of personhood, on which persons are to be *identified* with their self-network, it seems that there will inevitably be a chancy element to choice that cannot be attributed to the person, such that the kind of control that is exercised is too weak to ground his responsibility for which of the causal possibilities is realized.²⁸

However, I have not yet exhausted all of Kane's moves on this score. He considers the objection that invites us to consider an imaginary scenario in which two persons have exactly the same past and engage in identical efforts of will, but one of them chooses prudently while the other does not. Doesn't it seem a matter of luck, the objection goes, that one of them overcame short-sighted temptation, while the other did not? To connect this to the way I have been formulating the problem, we could say that it seems that the *deepest* explanation of the different outcomes is in terms of the causal potentialities of the properties constituting the agents' (identical) states prior to their choices, as characterized by the relevant statistical laws. If we had several such agents, a certain predictable proportion of them would choose one way and the rest would choose another. There is nothing more to be said.

To this, Kane replies that

[w]ith indeterminate efforts exact sameness is not defined. Nor is exact difference either. If the efforts are indeterminate, one

cannot say the efforts had exactly the same strength, or that one was exactly greater or less great than the other. That is what indeterminacy amounts to. So one cannot say of two agents that they had exactly the same pasts and made exactly the same efforts and one got lucky while the other did not.²⁹

He goes on to say that

[t]his is how free will is related to the uniqueness of persons. . . . Each life history is unique and cannot be exactly the same as any other if the psychological histories involve indeterminate processes, as they must for free will. . . .³⁰

An Epicurean world in which undetermined events occurred given an entirely determinate past—a world of chance without indeterminacy—would be a world of mere chance, not free will. There would be no indeterminate “gestation period” for free acts, so to speak; they would just pop out of a determinate past one way or the other without any preparation in the form of indeterminacy-producing tension, struggle, and conflict.³¹

It’s not clear to me why Kane claims (in the last sentence quoted) that there could be deliberative struggle only in a world of (amplified) indeterminacy. But, in any case, Kane’s basic idea here seems to be that genuine indeterminacy of states renders illegitimate any talk of exact sameness or difference. But if so, then it would likewise be illegitimate to talk of statistical laws that govern *types* of states and their range of possible outcomes, and so my way of raising doubts about the adequacy of Kane’s account of agent control could not get off the ground. However, this claim seems to me mistaken. For can we not speak of sameness of *superposition* of states—i.e., states having the same properties within the same *value intervals*? It seems to me that this is precisely how one would want to describe quantum mechanical states, and it renders perfectly meaningful statistical laws that quantify over such state types. Consequently, the causal indeterminist cannot exploit causal *indeterminacy*, in the way Kane has suggested, in order to avert the original problem of how one might control the “chancey” outcome of one’s reason states and “self-network.”

V. CONCLUSION

Progress in philosophical understanding does not always come by jettisoning seemingly baroque traditional accounts in favor of newfangled ones. Such is the case with the ancient conundrum of free will. There is much work to be done in explicating the picture of agency that centers on the notion of agent causation. (Let alone the work required to understand how to fit it into

a recognizable picture of the natural world more generally.) But the foregoing suggests that a satisfying understanding of free will cannot be had in any other way.³²

NOTES

1. That these reasons are familiar does not, of course, imply that they are obvious or uncontroversial.
2. Many of these accounts are included in Timothy O'Connor, ed., *Agents, Causes, and Events: Essays on Indeterminism and Free Will* (New York: Oxford University Press, 1995), sec. 2.
3. Timothy O'Connor, "Indeterminism and Free Agency: Three Recent Views," *Philosophy and Phenomenological Research* 53 (1993): 499–526.
4. Timothy O'Connor, "Agent Causation," in *Agents, Causes, and Events*, and "The Metaphysics of Free Will" (manuscript).
5. For contemporary defenses of such a view, see the contributions of Galen Strawson and Thomas Nagel to *Agents, Causes, and Events*.
6. As I emphasize below in the context of assessing Robert Kane's causal indeterminist view, the *kind* of control that agent causation is intended to capture is a rational, free control. More analysis is needed, though, to make plain the way in which agent-causal capacity is bound up with conscious awareness and rationality.
7. Carl Ginet, *On Action* (Cambridge: Cambridge University Press, 1990), 15.
8. *Ibid.*, 13.
9. *Ibid.*, 23ff.
10. *Ibid.*, 32.
11. See Carl Ginet, "Freedom, Responsibility, and Agency," *The Journal of Ethics* 1 (1997): 85–98.
12. I note that another contemporary simple indeterminist, Stewart Goetz describes (in unpublished work) his notion of a causally simple "choice" as the "exercising of mental power." But such a description is clearly inappropriate for a causally simple occurrence.
13. In chapter 2 of *Explaining Behavior* (Cambridge, Mass.: MIT Press, 1988), Fred Dretske notes that there are actually two senses to be given to the idea of something's causing a complex process of the form *F's causing G*. Consider the example of my pressing your doorbell (*F*) resulting in the ringing of the bell (*G*). Had the button not been wired to the bell in the appropriate way, *G* would not have obtained. The electrician who installed the device, it seems, was in some sense causally responsible for *F's causing G*, even though he did not in any way initiate this event. In Dretske's terminology, the electrician's activity was the *structuring cause* of *F's causing G* (by setting up the circumstances in such a way that *F*, should it occur, would result in *G*), though not its *triggering cause* (the cause of *F*, my pushing the button). But this distinction is not relevant to Ginet's claim, which clearly concerns the idea of an event's being the triggering cause of an agent-causal event.
14. And not his only, since many others have made this suggestion to me in conversation.
15. This charge is also leveled (in a very different fashion) by philosophers who reject simple indeterminism.
16. See, for example, Stewart Goetz, "A Noncausal Theory of Agency," *Philosophy and Phenomenological Research* 49 (1988): 303–16.
17. Ginet discusses this sort of example in the context of giving the objection to agent causation that is presently being discussed, but he draws a different conclusion, which is connected to his ontological commitment to finely individuated contrastive events.

18. Robert Kane, *Free Will and Values* (Albany, N.Y.: SUNY Press, 1985).
19. O'Connor, "Indeterminism and Free Agency: Three Recent Views."
20. Robert Kane, *The Significance of Free Will* (New York: Oxford University Press, 1996).
21. *Ibid.*, 23.
22. *Ibid.*, 125.
23. *Ibid.*, 128.
24. *Ibid.*, 129–30.
25. *Ibid.*, 127.
26. *Ibid.*, 139–40.
27. *Ibid.*, 140.
28. The picture seems even more inadequate for our ordinary conception of self-determination if it is combined with an ontologically reductionist account of token mental states (which Kane himself seems inclined to hold). For if the states that constitute the self-network have a microphysical complexity, such that the self-network's causal propensity (to give rise to any of a range of choice outcomes) is a direct function of the total set of individual propensities of each microevent (as described by fundamental statistical laws) to cause any of a range of other microevents, it is patently clear that which choice the network will indeterministically cause to be formed is not something that I myself directly control.
29. Kane, *The Significance of Free Will*, 171–72.
30. *Ibid.*, 172.
31. *Ibid.*, 173.
32. A shortened version of this paper was read at the 1996 Pacific Division Meetings of the Society of Christian Philosophers. I would like to thank that audience for helpful discussion, especially Michael Murray, who pressed me then and in subsequent correspondence on a couple of central claims I make above. The penultimate draft was also discussed at the 1996 Midwest Conference on Metaphysics at the University of Notre Dame. The probing comments by several members of the audience have helped me to clarify the issue (as I see it) between the agent-causal and the causal indeterminist view. I thank in particular Peter van Inwagen and Ted Warfield. I have also benefited from detailed written comments from James Felt.