Journal of Philosophical Theological Research
Print: 1735-9791
Online: 2538-2500

Making Sense of a Free Will that is Incompatible with Determinism: A Fourth Way Forward

Robert Kane*

Received: 2021-08-11 | Revision: 2021-08-15 | Accepted: 2021-08-23

Abstract

For a half - century, I have been developing a view of free will that is incompatible with determinism and, in the process, attempting to answer the Intelligibility Question about such a free will: Can one make sense of an incompatibilist or libertarian free will without reducing it to mere chance, or mystery, and can such a free will be reconciled with modern views of the cosmos and human beings? In this paper, I discuss recent refinements to my earlier writings on such a view, refinements developed in recent years in response to the large critical literature on my views in the past several decades. My view has usually been designated an event-causal (EC) view of libertarian free will and distinguished from non-causal (NC) and agent-causal (AC) libertarian views. But I was never happy with this designation of my view as "event-causal" and did not use it myself in earlier writings. In this paper, I explain why I now reject it altogether. I have come to believe that to avoid numerous misunderstandings in current debates about free will, we must distinguish four different kinds of libertarian theories, not merely three: in addition to non-causal (NC), agent-causal (AC), and event-causal (EC) theories, we need to add a fourth kind, which might be called an agent-causal/event-causal (AC/EC) theory. My view has always been of this fourth kind. It represents what I call in the title of this paper the "fourth way forward" for making sense of an incompatibilist free will.

Research Article

Keywords

Free will, Incompatibilism, Libertarianism, Agent-causal, Event-causal, Non-causal, Determinism.

* Distinguished Professor of Philosophy, Department of Philosophy, University of Texas, Austin. USA. robertkane@utexas.edu

▣ Kane, R. (2021). Making Sense of a Free Will that is Incompatible with Determinism: A Fourth Way Forward. *Journal of Philosophical Theological Research*, *(special issue on Free Will)*, 23(89), 5-28.
doi: 10.22091/jptr.2021.7247.2595

1. Introduction

For those who wish to defend an incompatibilist or libertarian free will requiring ultimate responsibility, one task they must confront is to offer arguments attempting to show that such a free will is not compatible with determinism. A second task is to address questions about whether an incompatibilist free will requiring ultimate responsibility is *intelligible* or *possible* at all and whether it can be reconciled with modern views of humans and the cosmos. This second question is what I have called the "Intelligibility Question" about an incompatibilist or libertarian free will and it is in many ways even more difficult than the Compatibility Question which is about whether free will is or is not compatible with determinism.

The culprit in the case of this second question is not determinism, but indeterminism. The Intelligibility Question is related to an ancient dilemma: If free will is not compatible with determinism, it does not seem to be compatible with indeterminism either. Arguments have been made since ancient times that undetermined events would occur spontaneously and hence could not be controlled by agents in the way that free and responsible actions would require. Undetermined events occurring in brains or bodies, it has long been argued, would not seem to enhance freedom and control over, and hence responsibility for, actions, but rather to diminish freedom, control, and responsibility. Arguments such as these and many others have led to often-repeated charges that undetermined choices or actions would be "arbitrary," "random," "irrational," "uncontrolled," "mere matters of luck" or "chance," and hence not free and responsible actions at all.

It is little wonder that libertarians about free will, those who believe it must be undetermined, have often looked for some deus ex machina to solve these problems, while their opponents have cried magic or mystery. Indeterminism might provide causal gaps in nature, libertarians have reasoned, but that was only a negative condition for free will. Some additional form of agency or causation was needed that went beyond familiar modes of causation in the natural order to "fill" those causal gaps in nature. And thus we had historical appeals to "extra factors," such as noumenal selves outside space and time (Immanuel Kant), immaterial minds (Rene Descartes), uncaused causes, and nonevent agent causes, to account for an otherwise undetermined free will. Tempting ways to think, perhaps, but also prompting charges such as P. F. Strawson's in his seminal 1962 paper "Freedom and Resentment" about the "panicky metaphysics" of libertarian free will.

I long ago became disenchanted with all such appeals to extra factors such as noumenal selves and the rest to account for an incompatibilist free will.

Avoiding them, however, requires a whole new look at the Intelligibility Question which is my goal in this paper: Can one make sense of a libertarian or incompatibilist free will without reducing it to mere chance, on the one hand, *or to mystery*, on the other, and can such a free will be reconciled with modern views of the cosmos and human beings?

2. Indeterminism: Empirical and Philosophical Questions

First, let us be clear that it is an empirical and scientific question whether any indeterminism is there in nature in ways appropriate for free will—in the brain, for example. No purely philosophical theory can settle the matter. As the Epicurean philosophers said centuries ago, if the atoms don't "swerve" in undetermined ways, and in the right places, there would be no room in nature for free will.

Christoph Koch is a distinguished neuroscientist and a tough-minded one at that. He argues that "there is no evidence that any components of the nervous system—a warm and wet tissue strongly coupled to its environment—display quantum entanglement." (2009, p.40) But Koch goes on to say that "what cannot be ruled out is that tiny quantum fluctuations deep in the brain are amplified by deterministic chaos" so that they might have non-negligible nondetermined effects on neural processing and thereby affect human decision-making (Koch, 2009, p.40) Koch does not endorse this idea, but says that it cannot be ruled out, given what is currently known about the brain. And such a role for indeterminism is all that would be needed for the theory to be presented here.

In the most recent edition of *The Oxford Handbook of Free Will* (2011), Robert Bishop agrees with Koch and cites a number of other neuroscientists and philosophers who have made similar suggestions. If minute quantum indeterminacies occurred at the neural or synaptic levels of the brain, Bishop argues, affecting the timing of the firing of individual neurons, these indeterminacies, however minute, could be amplified, due to sensitivity to initial conditions, so that they had non-negligible effects on neural processing. Bishop goes on to point out that one need not even have to appeal only to chaos to get these effects. For, as he notes, "the exquisite sensitivity needed for the normal amplification of quantum effects is a general feature of nonlinear dynamics" (Bishop, 2011, p.91). And it is generally agreed, Bishop adds, that nonlinear dynamics is pervasive in the functioning of human brains. Other scientists, not mentioned by Bishop, have also made suggestions about the possible role of indeterminism in the brain in recent years, including,

interestingly, its potential evolutionary significance.[1]

But our question at present is a philosophical one that has boggled people's minds since the time of the ancient Stoics and Epicureans: What could one *do* with indeterminism, assuming it was there in nature in the right places, to make sense of free will as something other than mere chance or randomness and without appealing to mystery? If minute quantum indeterminacies in the firings of individual neurons were amplified so that they introduced some indeterminism into the larger scale mental processing in deliberation and decision-making, how could that help to make sense of free choice as something other than mere chance? I will now try to explain how this might be.

3. Initial Pieces: Self-formation, Efforts, Will-power, Volitional Streams

Before continuing, however, let me add that those familiar with my past work will notice refinements of my earlier writings in the theory that follows, refinements that have been developed in recent years in response to the critical literature. In this regard, I owe a debt to many critics and other writers on these subjects. In the spirit of Plato's intimation that the search for Truth (with a capital T, one of his ideal forms, like Justice and the Good) is not something any one person or group can own and hoard from others, like a pot of gold. The Truth in this ideal sense is rather something we can only "participate in" (in Plato's words) with others to some degree from our own necessarily limited perspectives.

I will begin by reiterating a long-time theme in my writings, namely that indeterminism does not have to be involved in all actions done "of our own free wills." Indeterminism need be involved only in those choices or acts by which *we make ourselves into the kinds of persons we are, with the wills we have*. These seminal choices or acts in our lives are what I call "self-forming actions" or SFAs.

I argue that these self-forming actions occur at those difficult times in life when we are torn between competing visions of what we should do or

1. They include, among others, neuroscientist, Peter Ulric Tse, who has made detailed and highly original suggestions about these topics in a recent book (2013), as well as neuroscientists Paul Glimcher (2005), Michael Shadlin (2014), Biologists Bjorn Brembs (2010), Martin Heisenberg (2013), Astrophysicist Robert Doyle (2011), Physicists G. F. R. Ellis (2009), John Polkinghorne (2009), Hans Briegel (with philosopher Thomas Mueller) 2015, Psychologist Dean Simonton (2004), Among others (see, e.g., the survey of Jedlicka, 2014).

become, and they are more common in everyday life than one may think. Perhaps we are torn between doing the moral thing or acting from ambition, or between present desires and long term goals, or we are faced with difficult tasks for which we have aversions, or have to exert will-power to keep prior commitments and resolutions rather than break them.[1] In all such cases and many others, we may be faced with competing motivations and have to strive or make an effort or exert will-power to overcome the temptation to do something else we also strongly want.

At such times, the tension and uncertainty we feel about what to do, I suggest, would be reflected in some indeterminacy in our neural processes themselves. This would be in the form of amplified background neural indeterminacy as described in the previous section *"stirred up,"* one might say, *by the conflicts in our wills*. The uncertainty and inner tension we feel at such soul-searching moments of self-formation would thereby be reflected in some indeterminacy in our neural processes themselves. The experienced uncertainty would correspond physically to the opening of a window of opportunity that temporarily screens off complete determination by the past.

A further step then involves noting that in most such cases of self-formation, where we are faced with competing motivations, whichever choice is made will require an effort of will or exercise of will-power to overcome the temptations to make an alternative choice. I thus postulate, in such cases, that different goal-directed cognitive processes (*volitional streams*, we might call them) would be involved in the brain, corresponding to these exertions of effort or will-power. These cognitive processes would have different goals corresponding to the different choices that might be made (e.g., a moral choice or a self-interested choice).

Importantly, however, it is not being claimed here (as I have done in earlier writings) that these efforts or exercises of will-power aimed at different choices would be occurring at the same time during deliberation. Nor will they be occurring throughout the entire deliberation. Rather, different efforts or exertions of will-power may be initiated at different times, depending on the course of the agent's reasoning or thought processes.

To illustrate, consider a familiar example of Peter van Inwagen's of a would-be thief, call him John, deliberating about whether or not to steal from a church poor box. Suppose John is deeply torn because, on the one hand, he is desperately in need of money and knows that no one is usually in the church

1. On the importance of will-power and exercises of it, which may also involve effort making, I have learned much from Richard Holton's work (2009).

on weekday afternoons, so he can likely steal without being caught. On the other hand, he has serious moral qualms about doing so, made even stronger because he knows the money in the poor box is used to help other people who are in need, some of whom may need it as much as, or more than, he does.

We might then imagine that in the course of John's deliberation, various thoughts, experiences, and memories come to mind, various thoughts, desires, and possibilities are assessed, so that his considered reasons incline him to choose to steal the money rather than not to steal it. Of crucial importance, however, if this is to be a self-forming action or SFA in the sense understood here, we must say that the reasons motivating the choice to steal the money merely *incline* John to make that choice at this time rather than the alternative. These reasons are not "decisive" or "conclusive," nor do they determine he will do so. To use a traditional expression of Leibniz, his reasons "incline without necessitating." If a choice is thus to be made in accord with these inclinations, efforts would have to be made or will-power exerted to overcome the still-existing resistance in his will. This resistance would be coming from his motives to make the contrary choice, to refrain from stealing for moral reasons, which motives also remain important to him.

Though the reasons merely incline and are not decisive or conclusive, nonetheless when it is the case that the agent judges that they incline to one choice (e.g., to steal the money out of dire need) to a degree that might justify making that choice at the time, an effort would be initiated to make that choice and thereby to overcome the still-existent resistant motives in the will. This is where indeterminism enters the picture as well. For the conflict in John's will in such a case would "stir up" indeterminism in the effort to make the choice to which he is currently inclined, making it uncertain the effort will succeed in attaining its goal. If the effort to choose to steal does succeed, despite this indeterminism, the choice to steal to which John is presently inclined would be made and the deliberation would terminate.

Note that if this should happen, the choice to steal would have been made by John purposefully and in accordance with his will since it would have been the result of a goal-directed cognitive process (the effort or exertion of will-power) to make just this choice at this time rather than any alternative. Moreover, the choice would have been made for the reasons inclining him towards that choice at the time rather than the alternative. Thus it wouldn't have been a mere accident that the choice occurred, *even though its occurrence was undetermined*. The choice would have been brought about *voluntarily* and *on purpose* (that is, *intentionally*) *for* these inclining reasons, as a result of the success of the goal-directed effort of the agent that might

have failed due to the indeterminism, but did not fail.

But what would happen if, due to the indeterminism involved, the effort to choose to steal from the poor box did *not* succeed at that time and the choice had not been made? Many critics of a free will requiring indeterminism assume that if a choice is undetermined, the agent would be able to make a different choice, for example, to steal or not to steal, given exactly the same deliberation leading up to the moment of choice, including exactly the same desires, beliefs, thoughts, and prior reasoning. These critics then argue that, given this assumption, it would follow that, if John failed in his effort to choose to steal from the poor box at this time, due to the indeterminism ingredient in the effort, he would instead have chosen *not* to steal from the poor box at this same time instead. And this is problematic, these critics commonly argue, given that his deliberation would have been exactly the same leading up to the different choices. What would explain the difference in choice?

But it is important to note that this commonly made assumption by critics of an undetermined free will is *not* made in the account of self-forming choices being given here. It is *not* assumed, nor need it be assumed, on this account that if a choice is undetermined, the agent might make different choices, for example, to steal or not to steal, given exactly the same deliberation, including exactly the same desires, beliefs, inclinations, and reasoning, leading up to the choice. All that follows on this account from the assumption that a self-forming choice or SFA is undetermined is that the effort to make it may succeed *or may fail* at a given time in overcoming the resistance in the will to make it. And from this, it does *not* follow that if the effort fails, an alternative choice would be made at that same time instead.

To the contrary, failure would rather be a signal to the agent not to choose too quickly in terms of the presently inclining reasons. Failure would say in effect: Think more about this. The resistant motives for the alternative choice (in this case, his moral qualms about stealing the money) still matter to you more than you may have supposed and they should not be dismissed too readily. These resistant motives are the causal source of the indeterminism in the effort to choose to steal in the first place, making it uncertain that the effort will succeed here and now. The stronger these resistant motives are, the greater the probability the effort may fail, due to the indeterminism to which the resistant motives give rise.

What is assumed if, due to the indeterminism, John fails in his effort to choose to steal at the time is not that he would have made the contrary choice, not to steal, but rather that no choice at all would have been made at that time. The deliberation would either continue until a potential reassessment of the

motivating reasons led to another later effort to make the choice to steal or a potential reassessment led to a later effort to make the choice not to steal. Or, the deliberation might terminate without any decision being made, if this is possible in the circumstances and the agent is so inclined.

Note that in any of these possible scenarios, if John does succeed at a time in an effort to make whatever choice he is inclined to make at that time, he will have brought about that choice and will have done so voluntarily and intentionally and for the motivating reasons that inclined him towards that choice at the time. He would thus be responsible for the choice thereby made. For he would have succeeded in an effort whose goal was to make that very choice *rather than* the alternative *for* those inclining reasons; and this would be the case even though the occurrence of the choice, if he succeeds in his effort despite the indeterminism, would have been undetermined.

Indeterminism would have been involved in the effort, but it would not be the cause of the choice if the effort succeeds. For the effort would have succeeded, *despite* the indeterminism and not *because* of it. The cause of the choice would have been the agent whose effort or exercise of will-power it was. Note also that the indeterminism that is ingredient in the agent's effort to make the choice to which the agent is then inclined is not an *accidental* feature of the situation. This indeterminism does not just "happen" to be present. Its presence is rather a causal consequence of the conflict in the agent's will and of the resistant motives that are a feature of this conflict—resistant motives of the kind that have to be overcome by effort if the choice is to be made.

The idea is thus to think of the indeterminism involved in self-forming choices, not as a cause *acting on its own*, but as an *ingredient* in larger *goal-directed* or *teleological* activities of the agent, in which the indeterminism functions as a hindrance or interfering element in the attainment of their goals. The choices that result would then be *achievements* brought about by the goal-directed activities (the efforts or exercises of will-power) of the agent, which might have failed, given the indeterminism ingredient in them, but one or another of which may succeed in attaining its goal, despite this indeterminism. Note finally that when indeterminism functions in this manner as an obstacle to the success of goal-directed activities, the *indeterminism does not preclude responsibility, if the activities succeed in attaining their goals*, despite the indeterminism. For if the agents voluntarily and intentionally bring about this particular choice for reasons that they then and there endorse, and they are not determined to do so by anyone or anything else, they would be responsible for doing so.

4. Agency, Complexity, Deviant Causation, and Disappearing Agents

As readers of this paper are likely to be well aware, numerous objections have been made to incompatibilist accounts of free will requiring indeterminism, including this account. I have attempted to address many of the most common of these objections in past writings and cannot obviously discuss them all in this paper.[1] But I will address several key objections in the remainder of the paper, the answers to which involve the distinctive refinements of the view presented here. The first set of such objections to be discussed in this section and the next will also allow me to explain what I mean in the title of this paper by saying it represents "*a fourth way forward*" toward making sense of an incompatibilist free will.

This set of objections begins with a question that has had a hypnotic effect on modern free will debates, reflecting deeply-rooted intuitions: Do we not have to postulate an additional kind of "agent-causation" over and above causation by states and events to fully capture libertarian free choices, given that such choices must be undetermined by prior states and events? There is a residual fear underlying questions of this kind that the causal role of the *agent* will somehow "disappear" from the scene if we describe the agent's capacities and their exercise, including free will, in terms of causation by states and events involving the agent alone. This fear is understandable. But I believe it is a fear that is ultimately misguided.

A continuing substance (such as an agent) does not absent the ontological stage because we describe its continuing existence - its life, if it is a living thing - including its capacities and their exercise, in terms of states, events, and processes involving it. One needs more reason than this to think that agents do not cause things, only events cause things. For my part, I should confess that I am a substance ontologist and indeed something of an Aristotelian when it comes to thinking about the nature of living things and the relation of mind to body. Human agents are continuing substances with both mental and physical properties. But there is nothing inconsistent in saying this and being a "teleological intelligibility" theorist about free will who thinks that the *lives* of agents, their capacities, and the *exercise* of those capacities, including free will, must be spelled out in terms of states, processes, and events involving the agents.

In sum, *one does not have to choose between agent causation and event causation in describing freedom of choice and action*. One can affirm both, as

1. Kane 1989, 1994, 1996, 1997, 1999, 2000, 2002, 2005, 2007, 2009, 2011, 2014.

I would. In the case of self-forming choices or SFAs, for example, it is true to say both that "the agent's deliberative activity, including his or her effort, caused or brought about the choice," and to say that "the agent caused or brought about the choice." Indeed the first claim *entails* the second. Such event descriptions are not meant to deny that agents cause their free choices and actions. Rather the event descriptions spell out in more detail *how* and *why* the agents did so.

Note however that if event-causal descriptions are to have these implications, the event causes they describe must be "agent-involving" in a special manner. Relevant to explaining this special manner is a peculiarly modern scientific way of understanding human agency and causation by agents that has roots in the ancient view of Aristotle just mentioned. Agents, according to this modern conception with ancient roots, are to be conceived as *information-responsive complex dynamical systems.* "An agent's causing an action" is to be understood as "an agent, *conceived as such an information-responsive complex dynamical system*, exercising *teleological guidance control*, over some of its own processes."[1]

Complex dynamical systems are understood in this context in the manner of "dynamical systems theory." Such systems (now known to be ubiquitous in nature and which include living things) are systems in which emergent capacities arise as a result of greater complexity. When the emergent capacities arise, the systems as a whole or various subsystems of them impose novel constraints on the behavior of their parts. Alicia Juarrero, whose informative book on the nature and significance of complex dynamical systems in the sciences, *Dynamics in Action* (Juarrero, 1999), calls these emergent novel constraints on the behavior of parts of the system "*context-sensitive constraints.*"[2]

Such complex dynamical systems exhibit what I am calling here *teleological guidance control* (TGC) when they tend through feedback loops and error correction mechanisms to converge on a goal (called an attractor) in

1. Oxford Handbook, 2011, 396 ff.
2. Juarrero's discussion of complex dynamical systems is developed further in insightful ways in *Did My Neurons Make Me Do It?Philosophical and Neurobiological Perspectives on Moral Responsibility and Free Will* (2007) by philosopher Nancey Murphy and neuro-psychologist Warren S. Brown. In chapter 7 of their book, Murphy and Brown discuss my previous view of free will, agreeing with aspects of it, but criticizing some other aspects. I believe the fourth way developed in this paper addresses legitimate criticisms of my earlier view that they and others have raised. But arguing for this in complete detail is a task for further writing.
   For similar suggestions regarding agency and complex systems, see: Frith (2009) & Newsome (2009).

the face of perturbations. Such control, as neuroscientist Marius Usher argues (2006), *is necessary for any voluntary activity* and he interprets it in terms of dynamic systems theory, as I would as well. Neuroscientists E. Miller and J. Cohen (2001) argue that such cognitive guidance control in human agents stems from the active maintenance of patterns of activity in the prefrontal cortex that represent goals and the means to achieve them. These patterns provide signals to other brain structures whose net effect, as Miller and Cohen describe it, is "to guide the flow of activity along neural pathways that establish the proper mappings between inputs, internal states, and outputs."

An important consequence of understanding the agent causation involved in free agency and free will in this way is that the causal role of the agent in intentional actions of the kind needed for free agency and will is not *reducible* to causation by mental states of the agent alone. That would leave out the added role of the agent, qua complex dynamical *system*, exercising teleological guidance control over the processes *causally linking* mental states and events to actions.

The significance of this requirement can be shown by considering examples of "deviant causation" by mental states, such as desires and beliefs, in debates about action and free will. Consider Donald Davidson's well-known example of such causal deviance, which was put forward as a problem for Davidson's own influential account of intentional action in terms of causation by beliefs and desires (1963). The example is of a mountain climber who lets go of his rope, allowing his companion to fall. The climber *desires* to save his own life and *believes* he can do so in the present situation by letting go of his rope. But in the example, he does not intentionally let go of the rope. Rather, this desire and belief so unnerve him when he thinks of their consequences, that they cause him to *accidentally* let go of the rope.

What is lacking in such examples of deviant causation is the agent, understood as a dynamical system, exercising teleological guidance control over the *manner* in which the mental states cause the resulting events (by "guiding the flow of activity along neural pathways that establish the proper mappings between inputs, internal states, and outputs" and being able to alter those pathways in response to new information). In the absence of this *systemic control* by the agent, qua information responsive dynamical system, over the manner in which the mental states cause the resulting events, the causation by mental states would be "deviant" and the outcomes would not be intentional actions of the agent, but merely accidental occurrences.

A further significant consequence of understanding causation of free actions in this way, as Usher points out, is that, while teleological guidance control (TGC) of the kind required for voluntary action is compatible with determinism,

it is also compatible with indeterminism. A complex dynamical system can exhibit teleological guidance control, tending through feedback loops and error correction to converge on a goal, even when, due to the presence of indeterminism, it is uncertain whether the goal will be attained. Such teleological guidance control is thus available to libertarian theories of free will as well.

To sum up, one does not have to choose between agent causation and event causation in accounting for free agency, libertarian or otherwise. You can, indeed you must, affirm both. And the agent causation involved is not reducible to event causation by mental states and events alone for the reasons given. There is thus no "disappearing agent" problem as well.

5. Agent Causal/Event Causal (AC/EC)

These thoughts lead to the further significant refinement in my view that is related, as mentioned earlier, to "the fourth way forward" to making sense of an incompatibilist free will referred to in the title of this paper. It has become customary in the past several decades to distinguish libertarian views of free will into three kinds: *noncausal* (NC), *agent-causal* (AC), and *event-causal* (EC) views. Noncausal libertarian views hold that free choices or volitions are intrinsically active doings that are uncaused events. Agent-causal views hold that free choices and actions are caused by agents in a special manner that for various reasons (which differ depending on the particular version of the view) cannot themselves be caused by events, either deterministically or indeterministically. Event causal views, by contrast, hold that free choices or actions are caused by events, but indeterministically.

In the 1960s when I began thinking about these issues, most libertarian views were either noncausalist (NC) or agent-causalist (AC), or some combination of the two. Those who held dualist views of mind and body usually had elements of both views: Free choices or actions were uncaused by physical events, yet were caused in some manner by agents, qua immaterial minds. All such views and historical versions of them are what prompted Strawson's complaint in 1962 about the "panicky metaphysics" of libertarian free will.

What have since come to be called event-causal (EC) views were not yet on the horizon in those days when I began. Some philosophers, such as David Wiggins (1973), Robert Nozick (1981), Richard Sorabji (1980), and others, later suggested the possibility of libertarian theories that did not appeal to panicky metaphysics and might somehow be reconciled with contemporary scientific views. But my own view, developed in the 1970s and subsequent

decades, was the first to attempt to develop such a view in considerable detail. Other libertarian views with similar goals have been subsequently developed in the past several decades.

All libertarian views with such goals thereafter came to be designated "event-causal" (EC) libertarian views to distinguish them from traditional noncausal (NC) and agent-causal (AC) views. And since my view was one of the first of this new kind, it was often identified and discussed as a prime example of such an "event-causal" (EC) view. Yet I was never happy with this designation and did not use it myself in earlier writings, including *The Significance of Free Will* (1996). In that work and elsewhere I preferred to call my view a "teleological intelligibility" (TI) view of free will, an expression suggested by Gary Watson.

But I now think it necessary to go one step further and repudiate the simple designation "event-causal" altogether for my view because of its misleading implications. For one thing, the designation "event-causal" misleadingly suggests (and has suggested to many) that one must choose between saying that event causation, on the one hand, or agent causation, on the other, is fundamental to making sense of libertarian free will; and I have never believed that. What I believe and have consistently argued is what I have said earlier in this paper, namely, that both agent and event causation are necessary to account for libertarian free will and neither is reducible to the other.

To reduce agent causation to causation by mental states and events, as noted earlier, would leave out the *systemic control* by the agent, qua information responsive dynamical system, over the *manner* in which the mental states cause the resulting events. Without this systemic control, the causation by mental states would be "deviant" and the outcomes would not be intentional actions of the agent, but merely accidental occurrences. On the other hand, as also noted earlier, to leave out references to the causal role of mental states and events and merely to say that agents caused their free actions would also not suffice. For the event descriptions spell out in crucial detail *how* and *why* the agents caused their free actions, information that would be left out if one just said the agents caused their free actions.

As a consequence, I have come to believe that to avoid numerous potential misunderstandings in current debates about free will, we should distinguish at least four different kinds of libertarian theories, not merely three: to non-causal (NC), agent-causal (AC), and event-causal (EC) theories, we need to add a fourth kind, which might be called an agent-causal/event-causal (AC/EC) theory. My view has always been of this fourth kind. It differs from noncausal (NC) views in not requiring that basic libertarian free actions be uncaused. It differs from event-causal (EC) views in rejecting claims that libertarian free

actions can be adequately explained merely by claiming they are indeterministically caused in appropriate ways by beliefs, desires, intentions, and other mental states of agents. I argue that to explain what these "appropriate ways" amount to, one must bring in references to agents, as complex dynamical systems, exercising a certain kind of teleological guidance control over some of their own processes.

Finally, this fourth kind of (AC/EC) view differs from traditional agent-causal (AC) views in not having any of the following features, one or more of which has been claimed for agent-causal views by their historical and recent defenders: (1) The view to be defended here does not involve a "sui generis" or "metaphysically primitive" notion of causation by an agent or substance. (2) Nor does it involve a notion of agent causation that is not also required by compatibilist accounts of free agency. In addition, (3) the AC/EC view defended here does not appeal to a notion of agent causation whose *exercise* does not essentially involve causation by states, processes, and events (hence AC/EC). Nor does it require (4) a special kind of causation by an agent that is in principle, or by its very nature, incapable of being itself caused by prior events, either deterministically or indeterministically. Nor does the view developed here require (5) a special kind of causation by an agent that is not subject to or governed by laws of nature, nor (6) one that involves appeals to immaterial substances (Descartes and other substance dualists), noumenal selves outside space and time (Kant), prime movers unmoved, or other examples of what Strawson has called the panicky metaphysics of libertarian free will.

For all these reasons, I now prefer to designate my view as this fourth kind of incompatibilist view, as an agent-causal/event-causal or AC/EC view, distinguishing it from the other standard three views in the literature, non-causal (NC), agent-causal (AC) and event-causal (EC).[1] It represents the "fourth way forward" to making sense of an incompatibilist or libertarian free will, referred to in the title of this paper. I believe it allows one to answer a variety of common objections often made to the other standard views in the literature. In the final sections of this paper, I will discuss some of the most common of these objections having to do with luck and chance.

6. The Explanatory Luck Objection: Authors, Stories, Value Experiments, and Liberum Arbitrium

One of the most common and powerful variants of objections concerning luck

1. This designation was first introduced in my writings in Kane 2011.

that have been made against nearly all libertarian theories during the past three decades has been called the "Explanatory Luck Objection." It is stated in the following way by Alfred Mele, one of its most astute and persistent defenders:

> If [as indeterminism would seem to imply] different free choices could emerge from the same past of an agent, there would seem to be no explanation for why one choice was made rather than another in terms of the total prior character, motives, and purposes of the agent. The difference in choice, i.e., the agent's choosing one thing rather than another, would therefore be just a matter of luck (1998, p.582).

This objection in various formulations is now so widely cited and affirmed by critics of libertarian views of free will that it is often referred to as "the" luck objection in the literature. And many philosophers assume it is decisive. I think they are mistaken. But I also think the objection has the power it has because it teaches us something important about free will.

Note that there is an obvious response to this luck objection in terms of the AC/EC view described in the preceding section as the fourth way forward for making sense of an incompatibilist free will. For in the case of undetermined self-forming choices, as understood on this AC/EC view, it is *not* correct to assume, as this explanatory luck objection does, that "different free choices could emerge from the same past of an agent." This is not true if it means the agent could make opposing choices—for example, to steal or not to steal—given exactly the same prior reasoning leading up to the moment of choice. All that follows, as argued earlier, from the fact that a self-forming choice is undetermined is that it might be made at a given time or might *fail* to be made at that time. And it does *not* follow if it fails, that the opposing choice—not to steal—would be made at that same time, given exactly the same reasoning leading up to the choice to steal. Moreover, this would be true whichever choice is successfully made in an undetermined self-forming choice situation.

Moreover, the following things would also be true *whichever choice should be made* in a self-forming choice situation on this AC/EC view: (a) The agent would have caused or brought about the choice made by succeeding in a teleologically guided cognitive process (an effort) to bring it about, thereby overcoming resistance in the will to doing so. (b) The agent would have knowingly made that choice *rather than* the alternative. (c) The agent would have had the power to bring about the choice made and would have successfully exercised that power when it was made. The power was not unlimited, since the effort through which it was exercised might have failed, due to the indeterminism ingredient in it. But if the effort succeeded in reaching its goal, despite this indeterminism, the agent's power to make the choice would have been successfully exercised. (d) The choice would then

have been made for reasons that inclined (without necessitating) the agent to make that choice at that time rather than the alternative - reasons that the agent then and there chose to act upon. (e) The agent would have brought about the choice rather than the alternative *voluntarily* (without being coerced against his or her will) and (f) the agent would have done so *intentionally* or on purpose, not merely accidentally, by succeeding in a teleologically guided effort aimed at bringing about that very choice rather than the alternative for the inclining reasons for that choice.

If saying the agent's choosing one thing rather than the other is "just a matter of luck," as this objection does, is meant to deny any of these things (a)-(f) about such self-forming choices, then saying that the outcome was just a matter of luck seems to be the wrong conclusion to draw. And if one were to say that "just a matter of luck" is meant to be consistent with all these things, this explanatory argument from luck would seem to lose much of its traction.

Ah, but not quite all traction; and this is where things get interesting. With powerful arguments in philosophy, it is not enough to show that their conclusions do not necessarily follow from their premises. One needs also to show why they seem to have such power. The luck objection in this popular form does not show that libertarian free choices must be "just a matter of luck," if that entails denying any of these claims (a)-(f). But this luck objection does show that there is something to the often-repeated charge that such self-forming choices would be *arbitrary* in a certain sense.

A residual arbitrariness seems to remain in all self-forming choices because the agents cannot in principle have *sufficient* or *overriding* ("*conclusive*" or "*decisive*") prior reasons for making one option and one set of reasons prevail over the other. Therein lies the truth in this explanatory luck objection: An undetermined free choice *cannot be completely explained by the entire past,* including past causes or reasons; and I think it is a truth that reveals something important about free will generally.

I have argued elsewhere that such arbitrariness relative to prior reasons tells us that every undetermined self-forming choice is the initiation of a novel pathway into the future, a "value experiment," as I have elsewhere called it, whose justification lies partly in that future and is not fully explained by the past (1996, pp.145-146). In making such a choice we say, in effect, "I am opting for this pathway. It is not required by my past reasons, but is consistent with my past reasons and is one branching pathway my life can now meaningfully take. Whether it is the right choice, only time will tell. Meanwhile, I am willing to take responsibility for it, one way or the other."

Of special interest here, as I have often noted, is that the term "arbitrary" comes from the Latin *arbitrium*, which means "judgment"—as in *liberum*

arbitrium voluntatis ("free judgment of the will"—the medieval designation for free will since the time of Augustine) (1985, p.112; 1996, p.145). Imagine a writer in the middle of a novel. The novel's heroine faces a crisis and the writer has not yet developed her character in sufficient detail to say exactly how she will act. The author makes a "judgment" about this that is not determined by the heroine's already formed past, which does not give unique direction. In this sense, the judgment (*arbitrium*) of how she will react is "arbitrary," but not entirely so. It had input from the heroine's fictional past and in turn, gave input to her projected future.

In a similar way, as I have often emphasized, agents who exercise free will are both authors of and characters in their own stories at once (1996, p.146; 2002, p.425; 2011, p.401). By virtue of "self-forming" judgments of the will (*arbitria voluntatis*) (i.e., self-forming actions), they are *"arbiters"* of their own lives*, "making themselves*" out of a past that, if they are truly free, does not limit their future pathways to one. If we should charge such agents with not having sufficient or conclusive prior reasons for choosing as they did, they might reply:

> "True enough. But I did have 'good enough' reasons for choosing as I did, which I'm willing to endorse and take responsibility for. If they were not sufficient or conclusive reasons, that's because, like the heroine of the novel, I was not a fully formed person before I chose (and I still am not, for that matter). *Like the author of the novel, I am in the process of writing an unfinished story and forming an unfinished character who, in my case, is myself"* (Kane, 2002, p.425).

7. Contrastive Explanations

Another set of commonly made objections against views of free will requiring indeterminism are closely related to the explanatory luck objection just discussed. These are objections concerning "contrastive explanations." A contrastive explanation is an explanation for why one thing occurred *rather than* another. In the case of free choices, it would be an explanation in terms of an agent's prior character, reasons, or motives for why the agent made one choice rather than another.

The objection, in this case, takes the following form: If a self-forming choice (e.g., between A and B) is undetermined up to the moment when it is made, it may be argued that there could be no adequate contrastive explanation for why this choice was made rather than the alternative choice. For the fact that the choice was undetermined would mean that either choice (A or B) might have occurred, given the totality of the agent's traits of character,

motives, and reasoning preceding the moment of choice. There thus could not be an explanation for why one choice was made rather than the other at that moment in terms of the totality of the agent's character, motives, and reasoning prior to the choice.

The first thing to be said in response to this familiar argument is similar to the first thing said in response to the explanatory luck objection: On this fourth kind of libertarian view, the AC/EC view, defended here, it would not be true to say of self-forming choices, as this contrastive explanation objection does, that either choice (A or B) might have occurred at a given time, "given the totality of the agent's traits of character, motives, and reasoning preceding that moment of choice." All that follows on this AC/EC view from the fact that a self-forming choice (e.g., the choice A) is *undetermined* at a given time is that it might be made at that time or might fail to be made at that time. It does *not* follow that if the choice (A) fails to be made at that time, the opposing choice—(B)—*would* be made *at that same time,* given exactly the same reasoning that led to the choice A.

But those who make this objection concerning contrastive explanation to views of free will requiring indeterminism usually have a further assumption in mind that also needs to be addressed. They often assume as well that for an explanation of a free choice to be adequately contrastive in the sense they require, the following would have to be the case: If making the choice that was made rather than any alternative was the *rational* or *reasonable* thing to do, given the totality of the agent's reasons or motives, then making any alternative choice during that same deliberation, given the totality of the agent's reasons or motives, would *not* have been rational or reasonable. If making an alternative choice during this same deliberation might also have been a rational or reasonable thing to do, we would not have an adequate contrastive explanation, in the sense these critics would require, for why one choice was made *rather than* the other in terms of the totality of the agent's reasons and motives at the time.

But if this is what contrastive explanations must involve, in the sense these critics require, there clearly could not be contrastive explanations of self-forming choices, as understood on this AC/EC view. For, in addition to being undetermined on this AC/EC view, self-forming choices must satisfy *plurality conditions* for free choice: the power to make them and the power to do otherwise (e.g., to make some alternative choice) *either way,* voluntarily, intentionally *and rationally*. And this rules out the requirement, made by these critics that any other choice that might have been made in the course of a deliberation, other than the choice actually made, would have been unreasonable or irrational.

Moreover, this feature of satisfying plurality conditions is not a defect of self-forming choices, according to this AC/EC view, but rather it is a consequence of their power. For it is precisely because agents have the power to make such choices *and* the power to do otherwise, voluntarily, intentionally, and rationally *either way*, as this AC/EC view requires, that makes it possible for such choices to be *will-setting* rather than *will-settled*. Will-*settled* choices are those in which the agent's will is already settled on making one choice rather than the other by prior reasons *before* the choice is made. *Will-setting* choices by contrast are those in which the agent's will is set *in one way rather than another* only at the moment when one choice rather than another is made and not before. The power to make such will-setting choices at some points in our lives is what makes it possible for us *to be makers or creators to some degree of our own wills* rather than to be always acting from wills already formed.

It is also important to emphasize, however, that while agents who make such will-setting or self-forming choices may not have conclusive or decisive reasons for making the choice that is made rather than any other at the time (since their reasons only "incline without necessitating"), such agents do nonetheless have reasons for choosing that are "good enough" to render the choices they do make reasonable and rational ones, given their total reason sets when they choose. Some formal decision theorists speak in this connection of "*satisficing reasons*"—reasons that are good enough to justify a choice or decision even though they are not sufficient to render any possible alternative choice or decision that might have been made in the course of this deliberation unreasonable or irrational. Reasons for will-setting or self-forming choices are satisficing reasons in this decision-theoretic sense.

Moreover, the fact that the reasons for self-forming choices are satisficing in this sense is related to something important about free will that was spelled out in the prior section in response to the explanatory luck objection. It is related to the fact that "every undetermined self-forming choice is the initiation of a novel pathway into the future, whose justification lies partly in that future and is not fully explained by the past." In making such a choice it is appropriate to say, "I am opting for this pathway. It is not *required* by my past reasons, but is consistent with my past reasons and is one branching pathway my life may now meaningfully take."

This "narrative" conception of self-formation, as we might call it, is nicely captured in an important recent book by John Doris *Talking to Ourselves: Reflections on Selfhood, Agency, and Responsibility* (2015). In a section of this book in which Doris talks about my views of agency and responsibility, he notes (p.162) that in my defense of libertarian free will, I "develop the

intriguing suggestion that ambivalence," about what one's true values are "and its resolution in action," is not contrary to responsible agency, but is essential to responsible agency. It is so, at least at some points in our lives, when we are torn between conflicting values and are making self-forming choices.

At such times, Doris says, when on my view, as he correctly describes it, we are engaged in self-formation, it is possible that more than one path into the future could represent our "true values," and it would be "up to us" which path we will take. We decide then and there which of our *possible* true values our actions will express. If we were never ambivalent in this way—ambivalent as I have sometimes put it—we could not be self-creating beings, since our choices and actions would always be expressing what we *already* were, the formed will we already had.

At this point in his book where Doris references these views of mine on conflicting values and ambivalence, he also discusses the example of Huckleberry Finn—an example that has played such a prominent role in contemporary philosophical writings on agency and ethics. On one telling of the Huckleberry Finn story, Doris says, "Huck held values favoring *both* the conventional course of action," that he should turn his friend and companion Jim, a black man who had escaped from slave owners, over to the authorities and, on the other hand, "the course [Huck] actually followed," of not turning Jim over. In sum, Doris says, "Huck's values *conflicted*… he suffered a kind of *ambivalence*" (2015, p.161).

That, I believe, is how the author Mark Twain himself tells the story. Huck is growing and developing as a self or agent. In deciding not to turn Jim in to the authorities, Huck is not merely *expressing* what sort of a self he already is; he is also *deciding* what sort of a self he is going to be, by deciding from among the conflicting values he has, which ones he will follow. He is thereby not merely engaged in self-*expression*, but in self-*formation*, of the kind I believe *freedom of will* and not mere *freedom of action* sometimes requires. Such conflicts in the will and their resolution or lack thereof (as my wife, a writer, continually reminded me), are the stuff of most great literature and drama, Huckleberry Finn, Madame Bovary, Hamlet, Anna Karenina, you name it.

References

Bishop, R. (2011). Chaos, Indeterminism and Free Will. In R, Kane (Ed.), *The Oxford Handbook of Free Will* (2th ed., 84-100). Oxford: Oxford University Press:

Brembs, B. (2010). Towards a Scientific Concept of Free Will as a Biological Trait. *Proceedings of the Royal Society B: Biological Sciences,* 278, 930-939. **doi:** 10.1098/rspb.2010.2325

Briegel, H. & Thomas M. (2015). A Chance for Attributable Agency. *Minds and Machines,* 25 (3), 261-279. **doi:** 10.1007/s11023-015-9381-y

Davidson, D. (1963). Actions, Reasons and Causes. *Journal of Philosophy,* 60, 685-700.

Doris, J. (2015). Talking to Ourselves: Reflection, Ignorance and Agency. Oxford: Oxford University Press.

Doyle, B. (2011). *Free Will: The Scandal of Philosophy*. Cambridge, Mass: I-Phi Press.

Ellis, G. F. R. (2009). In N. Murphy, G.F.R. Ellis & T. O'Connor (Eds.), *Top Down Causation and the Human Brain* (pp.63-82). Berlin: Springer Verlag.

Frith, Ch. (2009). Free Will and Top-down Control in the Brain. In N. Murphy, G.F.R. Ellis & T. O'Connor (Eds.). *Downwards Causation and the Neurobiology of Free Will (*pp.199-210). Berlin: Springer Verlag.

Glimcher, P. W. (2005). Indeterminacy in Brain and Behavior. *Annual Review of Psychology*, 56, 25-56. **doi:** 10.1146/annurev.psych.55.090902.141429

Jedlicka, P. (2014). Quantum Stochasticity and (the End of) Neurodeterminism. In A. Corradini & U. Meixner (Eds.), *Quantum Physics Meets the Philosophy of Mind* (pp.183-197). Berlin: De Gruyter.

Juarrero, A. (1999). Dynamics in Action: Intentional Behavior as a Complex System. Cambridge, Mass: MIT Press.

Heisenberg, M. (2013). The Origin of Freedom in Animal Behavior. In A. Suarez & P. Adams (Eds.), *Is Science Compatible with Free Will*? (pp.95-103). New York: Springer.

Holton, R. (2009). *Willing, Wanting, Waiting.* Oxford: Oxford University Press.

Kane, R. (1985). *Free Will and Values*. Albany, N.Y.: State University of New York Press.

Kane, R. (1989). Two Kinds of Incompatibilism. *Philosophy and Phenomenological Research,* 50(2), 219-254. **doi:** 10.2307/2107958

Kane, R. (1994). Free Will: The Elusive Ideal. *Philosophical Studies* (*Special Issue: Symposium on Freedom, Responsibility, and Determinism*), 75, 25-60. **doi:** 10.1007/bf00989880

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

Kane, R. (1997). Free will, Responsibility and Will-setting. *Philosophical Topics*, 24(2), 67-90. **doi:** 10.5840/philtopics199624210

Kane, R. (1999). Responsibility, Luck, and Chance: Reflections on Free Will and Indeterminism. *Journal of Philosophy*, 96(5), 217-240. **doi:** 10.2307/2564666

Kane, R. (2000). Precis of The Significance of Free Will and 'Responses to Bernard Berofsky, J. M. Fischer & G. Strawson.' *Philosophy and Phenomenological Research*, 60, 157-167. **doi:** 10.2307/2653432

Kane, R. (2002). Some Neglected Pathways in the Free Will Labyrinth. In R.Kane (Ed.), *The Oxford Handbook of Free Will* (406-437). Oxford: Oxford University Press

Kane, R. (2005). *A Contemporary Introduction to Free Will*. Oxford: Oxford University Press.

Kane, R. (2007). Libertarianism. In J. M. Fischer, R. Kane, D. Pereboom & M.Vargas (Eds.), *Four Views on Free Will* (pp.5-43). Malden, Mass.: Blackwell.

Kane, R. (2007). Responses to Fischer, Pereboom and Vargas. In J. M. Fischer, R. Kane, D. Pereboom & M.Vargas (Eds.), *Four Views on Free Will* (166-183). Malden, Mass.: Blackwell.

Kane, R. (2009). Free Will and the Dialectic of Selfhood. *Ideas Valories*, 58, 25-44.

Kane, R. (2011). Rethinking Free Will: New Perspectives on an Ancient Problem. In R, Kane (Ed.), *The Oxford Handbook of Free Will* (2th ed., 381-404). Oxford: Oxford University Press.

Kane, R. (2014). New Arguments on Debates on Libertarian Free Will: Responses to Contributors. In D. Palmer (Ed.), *Libertarian Free Will: Contemporary Debates (*172-214). Oxford: Oxford University Press.

Koch, Ch. (2009). *Free Will, Physics, Biology and the Brain.* In N. Murphy, G.F.R. Ellis & T. O'Connor (Eds.), (pp.31-52). Berlin: Springer Verlag.

Mele, A. (1998). Review of Robert Kane, *The Significance of Free Will. Journal of Philosophy*, 95(11), 581-584. **doi:** 10.5840/jphil1998951122

Miller, E. & Cohen. J. (2001). An Integrative Theory of Pre-frontal Cortex Function. *Annual Review of Neuroscience*, 24, 167-202. **doi:** 10.1146/annurev.neuro.24.1.167

Murphy, N. & Brown, W. S. (2007). Did My Neurons Make Me Do It? Philosophical and Neuro-biological Perspectives on Moral Responsibility

and Free Will. Oxford: Oxford University Press.

Murphy, N. Ellis. G. F. R. & O'Connor, T. (Eds.) (2009). *Downwards Causation and the Neurobiology of Free Will.* Berlin: Springer Verlag.

Newsome, W. T. (2009). Human Freedom and 'Emergence'. In N. Murphy, G.F.R. Ellis & T. O'Connor (Eds.). *Downwards Causation and the Neurobiology of Free Will* (pp.63-82). Berlin: Springer Verlag.

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

Polkinghorne, J. (2009). *Questions of Truth.* Westminster: John Knox Press.

Shadlin, M. (2014). Comments on Adina Roskies: Can Neuroscience Resolve Issues About Free Will. In W. S. Armstrong (Ed.), *Moral Psychology*, (vol.4): *Free Will and Moral Responsibility* (pp.139-150). Cambridge MA: MIT Press.

Simonton, D. (2004). Creativity in Science: Chance, Logic, Genius, and Zeitgeist. Cambridge: Cambridge University Press.

Sorabji, R. (1980). Necessity, Cause and Blame: Perspectives on Aristotle's Philosophy. Ithaca: Cornell University Press.

Strawson, P. F. (1962). Freedom and Resentment. *Proceedings of the British Academy*, 48, 187-211. **doi:** 10.1073/pnas.48.1.1

Usher, M. (2006). Control, Choice and the Convergence/Divergence Dynamics: A Compatibilistic Probabilistic Theory of Free Will. *Journal of Philosophy*, 103(4), 188-213. 188-213. **doi:** 10.5840/jphil2006103431

Wiggins, D. (1973). Towards a Reasonable Libertarianism. In T. Honderich (Ed.), *Essays on Freedom and Action (*pp.31-61). London: Routledge & Kegan Paul.