

# Free Will: New Directions for an Ancient Problem

Robert Kane (2003)  
(edited by RJ)

## 1. The Compatibility Question: AP and UR

In a number of writings over the past two decades, I have sought to answer four questions about free will: (1) Is it compatible (or incompatible) with determinism? (2) Why do we want it? (3) Can we make sense of a free will that is incompatible with determinism? (4) Can such a free will be reconciled with modern images of human beings in the natural and social sciences?<sup>1</sup> On all four questions, I have tried to point current debates about free will in new directions. In this essay, I discuss some of these new directions.

Consider question (1) – the so-called Compatibility Question – which has received most of the recent attention in free will debates. The first thing we learn from these debates is that if we formulate the Compatibility Question as in most textbook discussions of free will – “Is freedom compatible with determinism?” – the question is too simple and ill-formed. The reason is that there are many meanings of “freedom” and many of them are compatible with determinism. Even in a determined world, we would want to distinguish persons who are free from such things as physical restraint, addiction or neurosis, coercion, compulsion, covert control or political oppression from persons who are not free from these things; and we could allow that these freedoms would be preferable to their opposites *even in a determined world*.

I think those of us who believe that free will is incompatible with determinism – we incompatibilists and libertarians so-called – should simply concede this point to our compatibilist opponents. Many kinds of freedom worth wanting are indeed compatible with determinism. What we incompatibilists should be insisting upon instead is that there is *at least one* kind of freedom worth wanting that is incompatible with determinism. This significant further freedom, as I see it, is “free will,” which I define as “the power to be the ultimate creator and sustainer of one’s own ends or purposes.” To say this further freedom is

important is not to deny the importance of everyday compatibilist freedoms from coercion, compulsion, political oppression, and the like; it is only to say that human longings transcend them.

This is one shift in direction for the Compatibility Question that I insist upon. But there is another of more importance. Most recent and past philosophical debate about the incompatibility of free will and determinism has focused on the question of whether determinism is compatible with “the condition of alternative possibilities” (which I shall call AP) – the requirement that the free agent “could have done otherwise.” Most arguments for incompatibilism, such as the “Consequence Argument” of van Inwagen and others, appeal to AP. Critics of such arguments either deny that AP conflicts with determinism or deny that alternative possibilities are required for moral responsibility or free will in the first place. As I view these contentious debates about AP and incompatibilism, they inevitably tend to stalemate over differing interpretations of “can,” “power,” “ability” and “could have done otherwise.” And I think there are good reasons for these stalemates having to do with the different meanings of freedom just mentioned. In response, I argue that we need to look in new directions. AP alone provides too thin a basis on which to rest the case for incompatibilism: *the Compatibility Question cannot be resolved by focusing on alternative possibilities alone*.

Fortunately, there is another place to look. In the long history of free will debate one can find another criterion fueling incompatibilist intuitions that is even more important than AP, though comparatively neglected. I call it the condition of ultimate responsibility or UR. The basic idea is this: to be ultimately responsible for an action, an agent must be responsible for anything that is a sufficient reason (condition, cause or motive) for the action’s occurring.<sup>2</sup> If, for example, a choice issues from, and can be sufficiently explained by, an agent’s character and motives (together with background conditions), then to be *ultimately* responsible for the choice, the agent must be at least in part responsible by virtue of choices or actions voluntarily performed in the past for having the character and motives he or she now has. Compare Aristotle’s claim that if a man is responsible for wicked acts that flow from his character, he must at some time in the past have been responsible for forming the wicked character from

which these acts flow.

This UR condition accounts for the “ultimate” in the original definition of free will: “the power of agents to be the *ultimate* creators and sustainers of their own ends or purposes.” Now UR does not require that we could have done otherwise (AP) for *every* act done of our own free wills – thus vindicating philosophers such as Frankfurt, Fischer, and Dennett, who insist that we can be held morally responsible for many acts even when we could not have done otherwise.<sup>3</sup> But the vindication is only partial. For UR *does* require that we could have done otherwise with respect to *some* acts in our past life histories by which we formed our present characters. I call these “self-forming actions,” or SFAs. Consider Dennett’s example of Martin Luther. When Luther finally broke with the Church at Rome, he said “Here I stand, I can do no other.” Suppose, says Dennett, at that moment Luther was literally right. Given his character and motives, he could not then and there have done otherwise. Does this mean he was not morally responsible, not subject to praise or blame, for his act, or that he was not acting of his own free will? Dennett says “not at all.” In saying “I can do no other,” Luther was not disowning responsibility for his act, but taking full responsibility for acting of his own free will. So “could have done otherwise,” or AP, says Dennett, is not required for moral responsibility or free will.

My response to Dennett is to grant that Luther could have been responsible for this act, even *ultimately* responsible in the sense of UR, though he could not have done otherwise then and there and even if his act was determined. But this would be so to the extent that he was responsible for his present motives and character by virtue of many earlier struggles and self-forming choices (SFAs) that brought him to this point where he could do no other. Those who know Luther’s biography know the inner struggles and turmoil he endured getting to that point. Often we act from a will already formed, but it is “our own free will” by virtue of the fact that we formed it by other choices or actions in the past (SFAs) for which we could have done otherwise. If this were not so, there is nothing we could have *ever* done to make ourselves different than we are – a consequence, I believe, that is incompatible with our being (at least to some degree) ultimately

responsible for what we are. So SFAs are only a subset of those acts in life for which we are ultimately responsible and which are done “of our own free will.” But if none of our acts were self-forming in this way, we would not be *ultimately* responsible for anything we did.

If the case for incompatibility cannot be made on AP alone, it can be made if UR is added; and thus, I suggest, the too-often neglected UR should be moved to center stage in free will debates. If agents must be responsible to some degree for anything that is a sufficient cause or motive for their actions, an impossible infinite regress of past actions would be required unless some actions in the agent’s life history (SFAs) did not have either sufficient causes or motives (and hence were undetermined). But this new route to incompatibility raises a host of further questions, including how actions lacking both sufficient causes and motives could themselves be free and responsible actions, and how, if at all, such actions could exist in the natural order where we humans live and have our being. These are versions of questions (3) and (4) listed above, which I call the Intelligibility and Existence questions for free will, to which I now turn.

## 2. The Intelligibility Question

The problem of intelligibility is an ancient one: if free will is not compatible with determinism, it does not seem to be compatible with indeterminism either. The arguments here are familiar and have been made since ancient times. An undetermined or chance event, it is said, occurs spontaneously and is not controlled by anything, hence not controlled by the agent. If, for example, a choice occurred by virtue of a quantum jump or other undetermined event in one’s brain it would seem a fluke or accident rather than a responsible choice. Or look at the problem in another way that goes a little deeper. If my choice is really undetermined, that means I could have made a different choice *given exactly the same past* right up to the moment when I did choose. That is what indeterminism and probability mean: exactly the same past, different possible outcomes. Imagine, for example, that I had been deliberating about where to spend my vacation, in Hawaii or

Colorado, and after much thought and deliberation had decided I preferred Hawaii and chose it. If the choice was undetermined, then exactly the same deliberation, the same thought processes, the same beliefs, desires and other motives – not a sliver of difference – that led up to my favoring and choosing Hawaii over Colorado, might by chance have issued in my choosing Colorado instead. That is very strange. If such a thing happened it would seem a fluke or accident, like that quantum jump in the brain just mentioned, not a rational choice. Since I had come to favor Hawaii and was about to choose it, when by chance I chose Colorado, I would wonder what went wrong and perhaps consult a neurologist. For reasons such as these, people have argued through the centuries that undetermined free choices would be “arbitrary,” “capricious,” “random,” “irrational,” “uncontrolled,” and “inexplicable,” not really free and responsible choices at all.

Defenders of an incompatibilist or libertarian free will have a dismal record of answering these familiar charges. Realizing that free will cannot merely be indeterminism or chance, they have appealed to various obscure or mysterious forms of agency or causation to make up the difference. Immanuel Kant said we can't explain free will in scientific and psychological terms, even though we require it for belief in morality.<sup>4</sup> To account for it we have to appeal to the agency of what he called a “noumenal self” outside space and time that could not be studied in scientific terms. Many other respectable philosophers continue to believe that only some sort of appeal to mind/body dualism can make sense of free will. Science might tell us there was indeterminacy or a place for causal gaps in the brain, but a non-material self, or what Nobel physiologist John Eccles calls a “transempirical power center,” would have to fill the causal gaps left by physical causes by intervening in the natural order.<sup>5</sup> The most popular appeal among philosophers today is to a special kind of *agent- or immanent causation* that cannot be explained in terms of the ordinary modes of causation in terms of events familiar to the sciences.<sup>6</sup> Free and responsible actions are not determined by prior events, but neither do they occur merely by chance. They are caused by agents in a way that transcends and cannot be explained in terms of ordinary modes of causation by events involving the agents.

I call these familiar libertarian strategies for making sense

of free will “extra factor” strategies. The idea behind them is that, since indeterminism leaves it open which way an agent will choose or act, some “extra” kind of causation or agency must be postulated over and above the natural flow of events to account for the agent's going one way or another. Early in my encounters with free will debates, I became disenchanted with all such extra factor strategies. I agree with other libertarian critics, such as Peter van Inwagen and Carl Ginet, that extra factor strategies – including agent-causal theories – do not solve the problems about indeterminism they are supposed to solve and they create further mysteries of their own.<sup>7</sup> If we are going to make progress on the Intelligibility and Existence questions about incompatibilist free will, I think we have to strike out in new directions, avoiding appeals to extra factor strategies altogether, including special forms of agent-causation. To do this means rethinking issues about indeterminism and responsibility from the ground up, a task to which I now turn.

### 3. Indeterminism and Responsibility

The first step is to note that indeterminism does not have to be involved in all acts done “of our own free wills” for which we are ultimately responsible, as argued earlier. Not all such acts have to be undetermined, but only those by which we made ourselves into the kinds of persons we are, namely “self-forming actions” or SFAs. Now I believe these undetermined self-forming actions or SFAs occur at those difficult times of life when we are torn between competing visions of what we should do or become. Perhaps we are torn between doing the moral thing or acting from ambition, or between powerful present desires and long term goals, or we are faced with a difficult task for which we have aversions. In all such cases, we are faced with competing motivations and have to make an effort to overcome temptation to do something else we also strongly want. There is tension and uncertainty in our minds about what to do at such times, I suggest, that is reflected in appropriate regions of our brains by movement away from thermodynamic equilibrium – in short, a kind of “stirring up of chaos” in the brain that makes it sensitive to micro-indeterminacies

at the neuronal level. The uncertainty and inner tension we feel at such soul-searching moments of self-formation is thus reflected in the indeterminacy of our neural processes themselves. What is experienced internally as uncertainty then corresponds physically to the opening of a window of opportunity that temporarily screens off complete determination by influences of the past. (By contrast, when we act from predominant motives or settled dispositions, the uncertainty or indeterminacy is muted. If it did play a role in such cases, it *would* be a mere nuisance or fluke, as critics of indeterminism contend.)

When we do decide under such conditions of uncertainty, the outcome is not determined because of the preceding indeterminacy – and yet it can be willed (and hence rational and voluntary) either way owing to the fact that in such self-formation, the agents' prior wills are divided by conflicting motives. Consider a businesswoman who faces such a conflict. She is on her way to an important meeting when she observes an assault taking place in an alley. An inner struggle ensues between her conscience, to stop and call for help, and her career ambitions which tell her she cannot miss this meeting. She has to make an effort of will to overcome the temptation to go on. If she overcomes this temptation, it will be the result of her effort, but if she fails, it will be because she did not *allow* her effort to succeed. And this is due to the fact that, while she willed to overcome temptation, she also willed to fail, for quite different and incommensurable reasons. When we, like the woman, decide in such circumstances, and the indeterminate efforts we are making become determinate choices, we *make* one set of competing reasons or motives prevail over the others then and there *by deciding*. ...

#### 4. Responsibility, Luck and Chance

You may find all this interesting and yet still find it hard to shake the intuition that if choices are undetermined, they *must* happen merely by chance – and so must be “random,” “capricious,” “uncontrolled,” “irrational,” and all the other things usually charged. Such intuitions are deeply ingrained. But if we are ever going to understand free will, I think will have to break old habits

of thought that support such intuitions and learn to think in new ways. The first step in doing this is to question the intuitive connection in most people's minds between “indeterminism's being involved in something” and “its happening merely as a matter of chance or luck.” “Chance” and “luck” are terms of ordinary language that carry the connotation of “its being out of my control.” So using them already begs certain questions, whereas ‘indeterminism’ is a technical term that merely precludes *deterministic* causation, though not causation altogether. Indeterminism is consistent with nondeterministic or probabilistic causation, where the outcome is not inevitable. It is therefore a mistake (alas, one of the most common in debates about free will) to assume that ‘undetermined’ means ‘uncaused.’

Here is another source of misunderstanding. Since the outcome of the businesswoman's effort (the choice) is undetermined up to the last minute, we may have the image of her first making an effort to overcome the temptation to go on to her meeting and then at the last instant “chance takes over” and decides the issue for her. But this is misleading. On the view I proposed, one cannot separate the indeterminism and the effort of will, so that *first* the effort occurs *followed* by chance or luck (or vice versa). One must think of the effort and the indeterminism as fused; the effort *is* indeterminate and the indeterminism is a property of the effort, not something separate that occurs after or before the effort. The fact that the effort has this property of being indeterminate does not make it any less the woman's *effort*. The complex recurrent neural network that realizes the effort in the brain is circulating impulses in feedback loops and there is some indeterminacy in these circulating impulses. But the whole process is her effort of will and it persists right up to the moment when the choice is made. There is no point at which the effort stops and chance “takes over.” She chooses as a result of the effort, even though she might have failed. ...

Perhaps the problem is that we are begging the question by assuming the outcomes of the woman's efforts are *choices* to begin with, if they are undetermined. One might argue this on the grounds that “if an event is undetermined, it must be something that merely *happens* and cannot be somebody's choice or action.” But to see how question-begging such a claim would be, one has

only to note what it implies: if something is a choice or action, it must be determined – that is, “all choices and actions are determined.” Is this supposed to be true of necessity or by definition? If so, the free will issue would be solved by fiat. ... Turning to choices, a choice is the formation of an intention or purpose to do something. It resolves uncertainty and indecision in the mind about what to do. Nothing in such a description implies that there could not be some indeterminism in the deliberation and neural processes of an agent preceding choice corresponding to the agent’s prior uncertainty about what to do. Recall from preceding arguments that the presence of indeterminism does not mean the outcome happened merely by chance and not by the agent’s effort. Self-forming choices are undetermined, but not uncaused. They are caused by the agent’s efforts.

Well, perhaps indeterminism does not undermine the idea that something is a *choice* simply, but rather that it is the *agent’s* choice. But again, why must it do that? What makes the woman’s choice her own on the above account is that it results from her efforts and deliberation which in turn are causally influenced by her reasons and her intentions (for example, her intention to resolve indecision in one way or another). And what makes these efforts, deliberation, reasons and intentions *hers* is that they are embedded in a larger motivational system realized in her brain in terms of which she defines herself as a practical reasoner and actor.<sup>8</sup> A choice is the agent’s when it is produced intentionally by efforts, deliberation and reasons that are part of this self-defining motivational system and when, in addition, the agent *endorses* the new intention or purpose created by the choice into that motivational system as a further purpose to guide *future* practical reasoning and action.

Well, then, perhaps the issue is not whether the undetermined SFA is a *choice*, or even whether it is the *agent’s* choice, but rather how much *control* she had over it. It may be true, as I argued earlier (in the discussion of plural voluntary control), that the presence of indeterminism need not eliminate control altogether. But would not the presence of indeterminism at least *diminish* the control persons have over their choices and other actions? ...

There is something to these claims, but I think what is true

in them reveals something important about free will. We should concede that indeterminism, wherever it occurs, *does* diminish control over what we are trying to do and *is* a hindrance or obstacle to the realization of our purposes. But recall that in the case of the businesswoman (and SFAs generally), the indeterminism that is admittedly diminishing her control over one thing she is trying to do (the moral act of helping the victim) *is coming from her own will* – from her desire and effort to do the opposite (go to her business meeting). And the indeterminism that is diminishing her control over the other thing she is trying to do (act selfishly and go to her meeting) is coming from her desire and effort to do the opposite (to be a moral person and act on moral reasons). So, in each case, the indeterminism *is* functioning as a hindrance or obstacle to her realizing one of her purposes – a hindrance or obstacle in the form of resistance within her will which has to be overcome by effort.

If there were no such hindrance – if there were no resistance in her will – she would indeed in a sense have “complete control” over one of her options. There would be no competing motives that would stand in the way of her choosing it. But then also she would not be free to rationally and voluntarily choose the other purpose because she would have no good competing reasons to do so. Thus, by *being* a hindrance to the realization of some of our purposes, indeterminism paradoxically opens up the genuine possibility of pursuing other purposes – of choosing or doing *otherwise* in accordance with, rather than against, our wills (voluntarily) and reasons (rationally). To be genuinely self-forming agents (creators of ourselves) – to have free will – there must at times in life be obstacles and hindrances in our wills of this sort that we must overcome.

Let me conclude with one final objection that is perhaps the most telling and has not yet been discussed. Even if one granted that persons, such as the businesswoman, could make genuine self-forming choices that were undetermined, isn’t there something to the charge that such choices would be *arbitrary*? A residual arbitrariness seems to remain in all self-forming choices since the agents cannot in principle have sufficient or overriding *prior* reasons for making one option and one set of reasons prevail over the other. There is some truth to this charge as well, but again I

think it is a truth that tells us something important about free will. It tells us that every undetermined self-forming free choice is the initiation of what I have elsewhere called a “value experiment” whose justification lies in the future and is not fully explained by past reasons. In making such a choice we say, in effect, “Let’s try this. It is not required by my past, but it is consistent with my past 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.”<sup>9</sup>

It is worth noting that the term “arbitrary” comes from the Latin *arbitrium*, which means “judgment”—as in *liberum arbitrium voluntatis*, “free judgment of the will” (the medieval philosophers’ designation for free will). 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, agents who exercise free will are both authors of and characters in their own stories all at once. By virtue of “self-forming” judgments of the will (*arbitria voluntatis*) (SFAs), they are “arbiters” of their own lives, “making themselves” out of past that, if they are truly free, does not limit their future pathways to one. ...

## 5. Agent Causation

When I began discussing the Intelligibility Question several sections ago, I said I would avoid appealing to any “extra factors” to account for libertarian free agency, such as noumenal selves, transempirical power centers, or special forms of agent- or nonevent causation, that libertarians have often appealed to. The preceding account makes no such appeals. It does appeal to the fact that free choices and actions can be caused by efforts, deliberations, beliefs, desires, intentions and other reasons or motives of agents. But this is causation by events or states of

affairs involving agents. It is not the special causation of agent-causal theories that cannot be spelled out in terms of events or states of affairs involving agents, either physical or psychological.<sup>10</sup> Moreover, causation by efforts, beliefs, desires, intentions and the like is something that even compatibilists appeal to in their accounts of free actions and choices; and it is hard to see how they could give accounts of free agency without doing so. The case is otherwise with such things as noumenal selves, transempirical power centers or nonevent causation, which are invoked specifically *to salvage libertarian intuitions* about free will and are not needed by non-libertarians.

This is what I mean by not invoking “extra” factors. My account of free will postulates no additional ontological entities or relations that non-libertarian accounts of free agency do not also need. It does postulate efforts, deliberations, desires, intentions and the like, and causation of actions by these. But compatibilists must postulate these also if they are going to talk about *free agency*. The only added assumption I have made to account for libertarian free agency is just what you would expect – that some of the mental events or processes involved must be *undetermined*, so that the causation by mental events may be nondeterministic or probabilistic as well as deterministic.

Of course, if any such theory is to succeed, there must be some indeterminism in the brain where undetermined efforts and choices occur. But such a requirement holds for any libertarian theory. If free choices are undetermined, as libertarians suppose, there must be some indeterminacy in the natural world to make room for them; and it is an empirical question whether the indeterminism is there. This is true even if one postulates special kinds of agent-causes or a non-material self to intervene in the brain. The indeterminism must be there to begin with in the brain, if these special forms of agency are to have room to operate. As the ancient Epicurean philosophers said, the atoms must sometimes “swerve” in undetermined ways, if there is to be room in nature for free will.

My suggestion about how indeterminism might enter the picture, if it were available in the physical world, was that conflicts in the wills of agents associated with self-forming choices would “stir up chaos” in the brain sensitizing it to quantum

indeterminacies at the neuronal level, which would then be magnified to effect neural networks as a whole. The brain would thus be stirred up by such conflict for the task of creative problem solving. This is speculative to be sure. Others have suggested different ways in which indeterminacy might be involved in the brain and free will.<sup>11</sup> But such speculations are not entirely idle either. There is growing evidence that chaos may play a role in human cognitive processing, as it does in many complex physical systems, providing some of the flexibility that the nervous system needs to adapt creatively to an ever-changing environment.<sup>12</sup> Of course, chaotic behavior, though unpredictable, is usually deterministic and does not of itself imply indeterminism. But chaos does involve “sensitivity to initial conditions.” Minute differences in the initial conditions of chaotic systems, including living things, may be magnified giving rise to large-scale undetermined effects. If the brain does “make chaos to understand the world” (as one recent research paper puts it<sup>13</sup>), its sensitivity to initial conditions may magnify quantum indeterminacies in neural networks whose outputs can depend on minute differences in the timing of firings of individual neurons. The general idea is that some combination of quantum physics and the new sciences of chaos and complexity in self-organizing systems may provide sufficient indeterminacy in nature for free will. What I emphasize is that only a small amount is needed in the precise timing of neuron firings. But this is only one idea among others. The question is ultimately an empirical one, to be decided by future research.

What I have tried to do in this paper is answer a different, but equally daunting, question: what could we *do* with the indeterminism to make sense of free will, supposing it were there in the brain? Wouldn't the indeterminism just amount to chance? How could it amount to free will unless one added some “extra factor” in the form of a special kind of agent-causation or transempirical power center to account for agency? As a final test of the answer given to these questions in this paper, it will be instructive to conclude with the following question: what is missing in the account of free will presented in earlier sections that an extra postulate of a special form of nonevent *agent-causation* is supposed to provide? We could ask the same question for other extra factor strategies, such as noumenal selves, transempirical

power centers, and the like. But most of these have gone out of favor in recent philosophy, while theories of nonevent agent- (or immanent) causation are still the most commonly discussed and defended libertarian theories today. So I will concentrate on contrasting agent-causal theories with the kind of libertarian theory I defend, which is often called *causal indeterminism*.

Let it be clear first of all that the causal indeterminist theory presented in this paper *does* postulate *agent causation* (though not of the nonevent or nonoccurrent kind). Agents *cause* or bring about their undetermined self-forming choices (SFAs) on this theory by making efforts to do so, voluntarily and intentionally; and agents cause or bring about many other things as well by making efforts to do so, such as deaths of prime ministers, broken tables, messes, accidents, fires, pains, and so on. Whether there is agent causation *in general* is not the issue here. What is at issue is agent-causation (hyphenated) – a *sui generis* form of causation postulated by agent-causal theorists that cannot be spelled out in terms of events and states of affairs involving the agents. It is misleading to frame this debate in such a way that libertarians who are agent-cause theorists believe in agent causation, while non-agent-causal libertarians like myself do not – presumably because we only believe in event causation. The fact is that both sides believe in agent causation. The issue is how it is to be spelled out.

And just as agents can be said to cause their self-forming choices (SFAs) and many other things, on the theory I proposed, so it can be said on this theory that agents *produce* or *bring about* their self-forming choices by making efforts to do so and *produce* many other things by their efforts and other actions. The point is worth making because defenders of agent-causation often claim that what causal indeterminist theories like mine lack – and what (nonevent) agent-causation is supposed to provide – is a conception of agents really *producing* or *bringing about* their undetermined free choices rather than those choices merely occurring by chance. But, as argued earlier, the mere presence of indeterminism does not imply that SFAs and other actions occur *merely* by chance and not as a result of the agent's voluntary and intentional efforts. Of course, the causation or production in the case of SFAs is nondeterministic or probabilistic, since they are

undetermined. And the burden of my argument was that such nondeterministic causation can support claims that agents really do *produce* what they cause by their voluntary efforts and can be held responsible for doing so.

So we are still looking for what the postulation of nonevent agent-causation is supposed to add to the picture that hasn't been captured. A perceptive recent defender of agent-causation, Timothy O'Connor, provides some further clues about this matter that are worth considering. Speaking to the issue of what causal indeterminist theories like mine lack that nonevent agent-causation is supposed to provide, O'Connor says the following.

[For causal indeterminist theories,] "the agent's internal states [including reasons, motives, etc.] have objective tendencies of some determinate measure to cause certain outcomes. While this provides an *opening* in which the agent might freely select one option from a plurality of real alternatives, it fails to introduce a causal capacity that fills it. And what better here than its being the agent himself that causes the particular action that is to be performed?"<sup>14</sup>

The missing element suggested in this quote is the "causal capacity" to "freely select one option from a plurality of real alternatives" that are left open by the (causal) indeterminism of prior events.

Now such a causal capacity is surely important. But why do we have to suppose that agent-causation of a nonevent kind is needed to capture it? The fact is that, on the causal indeterminist view presented, the agent *does* have such a causal capacity. Not only does the businesswoman facing an SFA have a plurality of real alternatives from which to choose, she has the *capacity* to make either choice by making an effort to do so. The conflicting motives in her will and the consequent divisions within her motivational system make it possible for her to choose either way for reasons, voluntarily and intentionally. And this is clearly a *causal* capacity since it is the capacity to *cause* or *produce* either choice outcome (nondeterministically, of course) as a result of her effort against resistance in her will.

This is a remarkable capacity to be sure; and we may

assume that it is possessed only by creatures who attain the status of *persons* capable of self-reflection and having the requisite conflicts within their wills. So O'Connor's calling it a form of "*personal* causation" is altogether apt. But there is no reason to suppose we need to postulate a *nonevent* form of causation to account for it. The *capacity* itself (prior to its exercise) is a complex dispositional *state* of the agent; and its *exercise* is a sequence of *events* or *processes* involving efforts leading to choice and formation of intention, which intention then guides subsequent action (of going back to help the victim or going on to a meeting.) This is a capacity *of* the agent, to be sure, but both the capacity and its exercise are described in terms of properties or states of the agent and in terms of states of affairs, events and processes involving the agent, as I have done in the preceding paragraph and earlier in the paper.

Is there a residual fear functioning here that the "agent" will somehow disappear from the scene if we describe its capacities and their exercise, including free will, in terms of states and events? Such a fear would be misguided at best. 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 of affairs, events and processes involving it. One needs more reason than this to think that there are no continuing things or substances, or no agents, but only events, or 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. Agents are continuing substances with both mental and physical properties. But there is nothing inconsistent in saying this and being a causal indeterminist 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 them.

Similar remarks are in order about O'Connor's comments about "emergence" or "emergent properties" of agents (such as emergent causal capacities) in connection with free will. Issues about the existence of emergent properties (like issues about continuing substances) must also be distinguished from issues



about nonevent causation. Indeed, I also believe that emergence of a certain kind (now recognized in self-organizing systems) is necessary for free will, even of the causal indeterminist kind that I defend. Once the brain reaches a certain level of complexity, so that there can be conflicts in the will of the kind required for SFAs, the larger motivational system of the brain stirs up chaos and indeterminacy in a part of itself which is the realization of a specific deliberation. In other words, the whole motivational system realized as a comprehensive “self-network” in the brain has the capacity to influence specific parts of itself (processes within it) in novel ways once a certain level of complexity of the whole is attained. This is a kind of emergence of new *capacities* and indeed even a kind of “downwards causation” (novel causal influences of an emergent whole on its parts) such as are now recognized in a number of scientific contexts involving self-organizing and ecological systems (Kuppers 1992; Kauffman 1995; Gilbert & Sarkar 2000).

But this kind of emergence characteristic of self-organizing systems does not, in and of itself, imply causation of a nonoccurrent or nonevent kind, since the wholes and parts involved are states and processes of the organism of various levels of complexity. Of course, O'Connor would like a stronger form of emergence, which would require nonoccurrent causation. But his argument – that some kind of emergence of capacities for holistic or downwards causation of wholes on parts is required for free will – does not prove the need for a *nonevent* kind of causation. Such emergence, which I agree is important for free will, can be accommodated within a theory of the kind I have proposed.

O'Connor offers yet another argument when he says that what non-agent-causal theories lack and what agent-causation supplies is “the agent's directly controlling the outcome” of an undetermined choice. This is the issue of *control* about which I have said a great deal earlier in this essay. What is it for an agent to have direct *control* at a given time over a set of choice options (e.g., to help the assault victim or go on to a meeting)? The answer given earlier is embodied in the idea of plural voluntary control. Stating it more precisely, agents have plural voluntary control over a set of options at a time when they have the (i) *ability* or *capacity* to (ii) *bring about* (iii) at that time (iv) *whichever* of the options

they will or want, (v) for the reasons they will to do so, (vi) on purpose or intentionally rather than accidentally, by mistake or merely by chance, hence (vii) voluntarily (in accordance with their wills rather than against them), (viii) as a result of their efforts, if effort should be required, (ix) without being coerced or compelled or (x) otherwise controlled or forced to choose one way or the other by some other agent or mechanism. Agents *exercise* such control *directly* when they voluntarily and intentionally *produce* one of the options (a particular self-forming choice or SFA) *then and there* (at the time in question) under these conditions. I have argued here and in other writings that these conditions can be satisfied for SFAs without appealing to any kind of nonevent agent-causation.<sup>15</sup> Moreover, these conditions of plural voluntary control are the kinds we look for when deciding whether persons are or are not *responsible* for their choices or actions (e.g., when they produce something voluntarily and intentionally as a result of making an effort to do so).

Finally, I want to consider an objection about control made to my theory by another agent-causal theorist, Randolph Clarke. Clarke argues that causal indeterminist theories, like mine, provide “leeway” for choice, but no more control over actions *than compatibilists offer*; and more control than compatibilists offer is needed to account for the genuine libertarian free will and responsibility.<sup>16</sup> I agree that something more in the way of control than compatibilists offer is needed to account for libertarian free will. But I think the “more” control libertarians need is not more of the same *kind* of control compatibilists offer, but rather another kind of control altogether. The kind of control that concerns compatibilists is what might be called “antecedent determining control” – the ability to guarantee or determine *beforehand* which of a number of options is going to occur. If free choices are undetermined, we cannot have antecedent determining control over them, for exercising such control would mean *predetermining* them – determining beforehand just which choice we are going to make. (Even nonevent agent-causation cannot give us that.) What libertarians must require for undetermined SFAs is I think another kind of control altogether (that compatibilists cannot get)--namely, *ultimate* control – the originaive control exercised by agents when it is “up to them” which of a set of possible choices or actions will

now occur, and up to no one and nothing else over which the agents themselves do not also have control. This is the kind of control required by ultimate responsibility or UR and it is not something that can be captured by compatibilists, since it requires indeterminism. But neither does such ultimate control require nonevent causation, as I have been arguing. What it does require the ability or capacity to cause or produce any one of a set of possible choices or actions each of which is undetermined (hence nondeterministically) and to do so “at will,” that is, rationally (for reasons), voluntarily and intentionally.

Note also that there is a trade-off between this ultimate control and the antecedent determining control that compatibilists want. To have ultimate control over our destinies, we have to give up some antecedent determining control at crucial points in our lives. We have to accept a measure of uncertainty and genuine indeterminacy right up to the moment of decision. Indeterminism does not leave everything unchanged, for it implies “the probability or chance of failure” – though with genuine free will, every failing is also a succeeding, so we are responsible either way. If libertarians were after the same kind of control that compatibilists have to offer – only more of it – then I would agree with Clarke. But I think that what motivates the need for incompatibilism is an interest in a different kind of “control over our lives” altogether – a control which has to do with our being to some degree the ultimate creators or originators of our own purposes or ends and hence ultimate “arbiters” of our own wills. We can’t have that in a determined world.

## NOTES

<sup>1</sup> See especially The Significance of Free Will (1996), which provides an overview of philosophical debates about all four questions over the past fifty years and further development of many of the ideas of this paper. Also, see an earlier work, Free Will and Values (1985) and the articles cited in the suggested reading after this essay and in the bibliography.

<sup>2</sup> For a formal statement and defense of this condition, see The Significance of Free Will, chapter 3.

<sup>3</sup> For defenses of this claim by these authors see the readings in this volume by Dennett, Fischer and Pereboom.

<sup>4</sup> Kant, The Critique of Practical Reason (1956), part III.

<sup>5</sup> Eccles, Facing Reality (1970).

<sup>6</sup> For discussion and defense of this view, see the readings in this volume by Chisholm and O'Connor and the suggested reading at the end of each of those readings.

<sup>7</sup> See the preceding essays in this Part by van Inwagen and Ginet.

<sup>8</sup> That some such motivational system is necessary to define personhood and agency has been persuasively argued by Fred Dretske (1988), David Velleman (1992) and Owen Flanagan (1992). In The Significance of Free Will (pp. 137-42), I call the realization of such a system in the brain, the “self-network.”

<sup>9</sup> The Significance of Free Will, pp. 145-6.

<sup>10</sup> I am aware that the nature of mental causation (or causation by mental states such as beliefs and desires) is itself a matter of controversy among philosophers. But I am making only two simple points about it here. First, since mental causation must be assumed by compatibilist accounts of free agency as well as libertarian accounts such as my own, whatever problems attach to the idea are not simply problems for libertarian theories or theories like mine. Second, causation by desires, beliefs, etc. is causation by states or events and does not commit one to nonevent agent-causation. I think both points are defensible. Some libertarians who are simple indeterminists, such as Ginet, would deny the first point (though not the second) since they argue that explanations of actions in terms of beliefs, desires and other mental states are not causal explanations at all. I disagree with this simple indeterminist view, but do not try to argue against it in this paper. See the preceding essay in this volume by Ginet and the suggested reading that follows it.

<sup>11</sup> See H. Stapp, Mind, Matter and Quantum Mechanics (1993), D. Hodgson, The Mind Matters (1991), John Eccles How the Self Controls the Brain (1994), R. Penrose, Shadows of the Mind (1994), I. Prigogine and I. Stengers, Order Out of Chaos (1984).

<sup>12</sup> H. Walter, Neurophilosophy and Free Will (2001), summarizes much of this recent research. See also C. Skarda and W. Freeman, “How the Brain Makes Chaos in Order to Understand the World” (1987) and A. Babloyantz and A. Destexhe, “Strange Attractors in the Human Cortex” (1985).

<sup>13</sup> Skarda and Freeman. (See note 12).

<sup>14</sup> O'Connor, “Libertarian Theories: Agent-causal and Dualist Theories” (forthcoming)

<sup>15</sup> See The Significance of Free Will, chapter 8 and “Responsibility. Luck and Chance: Reflections on Free Will and Indeterminism” (1999)

<sup>16</sup> Clarke and others have also posed questions about the (dual) “efforts of will” that precede self-forming choices or SFAs on my theory. The SFAs are nondeterministically caused by these preceding efforts, but are the efforts themselves determined by the agents prior reasons or motives? My answer is that the efforts agents make in SFA situations are causally influenced by their prior reasons or motives, but they are not strictly speaking determined by those reasons because the efforts themselves are indeterminate, which means there is

---

some indeterminism involved in the complex neural processes realizing them in the brain. Thus, the reasons do not determine that an exact amount of effort will be made. This means that indeterminism enters the picture in two stages, first, with the efforts, then with SFAs. One might say that, with the efforts, one opens a "window" of indeterminacy whose upshot is that the choice outcome (the SFA) will not be determined. But the primary locus of indeterminism is in the moment of choice itself, the SFA. The latter is undetermined in a way that allows for robust alternative possibilities (making a moral choice or an ambitious choice). To prepare for this, a measure of indeterminacy enters the picture earlier, in the preceding indeterminate efforts. A related question: do the agents cause these efforts? No, not in the way they cause their SFAs, because the efforts are basic actions. Agents make the efforts, they do not cause them by doing something else. And what it means to say they make the efforts was spelled out earlier (section 4) in the account of what it means to say that the businesswoman's choice was hers. Finally, are the efforts freely made? I distinguish three senses of freedom, all of which I think are required for a complete account of free action and free will: (i) not being coerced, compelled, controlled etc. (ii) acting "of one's own free will" in the sense of a will of one's own making (i.e., satisfying UR) and (iii) being an undetermined self-forming action or SFA. Sense (i) is compatibilist (and I think it is necessary for free will, though not sufficient); senses (ii) and (iii) are incompatibilist. Efforts of will preceding SFAs are free in senses (i) and usually (ii) also; SFAs (the full flowering of free will) are true in all three senses.