



## ORDER: GOD'S, MAN'S AND NATURE'S

### Rethinking Free Will: New Perspectives on an Ancient Problem

Robert Kane

#### Conference Paper: Where Free Will and Nature Meet

#### 1. Introduction

My dealings with free will date back to the mid-1960s and are coterminous with a resurgence of interest among philosophers in problems about the freedom of the will that began in the decade of the 1960s. The landscape of free will debate was simpler then. The unstated assumption was that if you had scientific leanings, you should be a compatibilist about free will (believing it to be compatible with determinism). That is, you should be a compatibilist, if you did not deny we had free will altogether (as did skeptics and hard determinists). And if you were a libertarian about free will—believing in a free will that is incompatible with determinism—you must (in order to make sense of such a free will) inevitably appeal to uncaused causes, immaterial minds, noumenal selves, non-event agent causes, prime movers unmoved, or other examples of what P. F. Strawson called the “panicky metaphysics” of libertarianism (in his influential 1962 essay “Freedom and Resentment”).

I started thinking about free will shortly after Strawson's essay appeared, when my philosophical mentor at the time, Wilfrid Sellars, challenged me to reconcile a traditional incompatibilist or libertarian free will with modern science. Sellars was a compatibilist about free will, like the vast majority of philosophers and scientists of that era; and like Strawson (whose essay he admired), he did not believe a traditional libertarian free will could be accounted for without appealing to obscure or mysterious forms of agency of the kinds

Strawson had dubbed "panicky metaphysics." Employing a well-known distinction he had introduced in the philosophical literature, Sellars granted that free will in some sense was an integral part of what he called "the manifest image" of humans and their world. But he did not believe a libertarian free will—one that was incompatible with determinism—could be reconciled with "the scientific image" of that world; and he challenged me to show otherwise.

I accepted the challenge at the time; and I remember thinking—with the brashness and naivete of youth: "Give me three or four weeks and I'll wrap this up and be back with an answer (or at the outside by the end of the semester!)" Well, it is now forty-five years later and I am still struggling with the challenge. One reason the challenge was more difficult than I naively assumed was something I only gradually came to realize: To make sense of a traditional free will of an incompatibilist or libertarian kind—which Nietzsche derisively called free will "in the superlative metaphysical sense"—one must learn to think in new ways, to break old molds of thought and substitute new ones. Otherwise I think such a freedom is likely to appear utterly mysterious, the "greatest self-contradiction" conceived by the mind of man, as Nietzsche went on to argue. In what follows, I discuss some of the new directions my thought had to take in order to make headway with the problem.

## **2. The Compatibility Question: Alternative Possibilities (AP)**

The first step was to take a new look at the Compatibility Question—whether free will compatible or incompatible with determinism—which has been at the center of current debates about free will over the past century.

The first thing we should learn from these debates, I came to believe, is that if the Compatibility Question is formulated 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" (as one would expect of such a protean and much-used term); and many of them are compatible with determinism. Even if we lived 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 should 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 about free will—should simply concede this point to our compatibilist opponents: Many kinds of freedom worth wanting are indeed compatible with determinism. What incompatibilists should insist 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 some 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 go beyond them.

This is one shift in direction for the Compatibility Question that I came to emphasize. But there is another of more importance. Most recent and past philosophical debate about the Compatibility Question has focused on the question of whether determinism is compatible with "the condition of alternative possibilities" (which I call AP)—the requirement that the free agent must have had alternative possibilities and hence the "power" or "ability" to have "done otherwise." Most arguments for the incompatibility of free will and determinism (of which the so-called "Consequence Argument" of Peter van Inwagen and others is the most well known) appeal to this AP condition in one way or another. These arguments claim that if determinism were true, agents could not have done otherwise, since only one alternative future would have been possible, given the past and laws of nature; and agents do not now have the power or ability to change either the past or the laws of nature. Compatibilist critics of such arguments have either denied that the power or ability to do otherwise (the AP condition) conflicts with determinism or have denied that being able to do otherwise is required for moral responsibility or free will in the first place.

As I view these contentious debates about alternative possibilities 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 to resolve the Compatibility Problem, we need to look in new directions. AP alone provides too thin a basis on which to rest the case for the incompatibility of free will and determinism: the Compatibility Problem cannot be resolved by focusing on alternative possibilities alone.

### **3. Ultimate Responsibility (UR) and Self-forming Actions (SFAs)**

Fortunately, there is another place to look for reasons why free will might conflict with determinism. I have argued that in the long history of free will debate, one can find a second criterion fueling incompatibilist intuitions even more important than AP, though comparatively neglected. I call it ultimate responsibility, or UR.<sup>i</sup> The 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. 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.<sup>ii</sup>

This UR condition does not require that we could have done otherwise (AP) for every act performed "of our own free wills"—thus partially vindicating those philosophers such as Frankfurt (1969), Dennett (1984), Fischer (1994) and others, who insist that we can be held morally responsible for many acts even when we could not have done otherwise. 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.<sup>iii</sup>

Consider Daniel Dennett's much-discussed example of Martin Luther (1984: 131-3). When finally breaking with the Church at Rome, Luther said "Here I stand, I can do no other." Suppose Luther was literally right about himself at that moment, says Dennett. Given his character and motives, he literally could not then have done otherwise. Does this mean he was not morally responsible for this act? Not at all, Dennett answers. In saying "I can do no other," Luther was not disowning responsibility for his act, but taking full responsibility for it; and thus "could have done otherwise," or AP, is not required for free will in a sense demanded by moral responsibility.

My response 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 even if his act was determined. But this would be so, I would argue, to the extent that Luther was responsible for his present motives and character by virtue of earlier struggles and self-forming choices (SFAs) that brought him to this point where he could do no other. 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 (which did satisfy AP). If this were not so, there would have been nothing we could have ever done in our entire lifetimes to make ourselves different than we are—a consequence, I believe, that is incompatible with being ultimately responsible (UR) for what we are.<sup>iv</sup> So, while SFAs are not the only acts in life for which we are ultimately responsible and which are done "of our own free will," 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 came to believe that 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 reason (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 sufficient causes or motives and hence were not determined.

Focusing on UR tells us something else of paramount importance about free will. It tells us why the free will issue is about the freedom of the will and not merely about freedom of action. There has been a tendency in the modern era of philosophy, beginning with Hobbes and Locke in the seventeenth century and coming to fruition in the twentieth century, to reduce the problem of free will to a problem of freedom of action. I have been arguing for some time that such a reduction oversimplifies the problem. Free will is not just about free action. It is about "self-formation," about the formation of our "wills," or how we got to be the kinds of persons we are, with the characters, motives and purposes we now have. Were we ultimately responsible to some degree for having the wills we do have, or can the sources of our wills be completely traced backwards to something over which we had no control— God, fate, heredity and environment, nature or upbringing, society or culture, behavioral engineers or hidden controllers? Therein, I believe, lies the core of the traditional "problem of free will."

#### **4. The Intelligibility Problem**

But this approach to the incompatibility of free will and determinism raises a host of further extremely difficult questions about free will. How, for example, could acts lacking both sufficient causes and motives be free and responsible actions? And how could such actions exist in the natural order where we humans live and exercise our freedom? Can we make sense of such a notion of free will or is it an unintelligible, impossible or self-contradictory ideal, as Nietzsche and many other modern philosophers and scientists believe? And can it be reconciled with modern scientific conceptions of humans and the cosmos?

Doubts about the very possibility of an incompatibilist free will are related to an ancient dilemma: if free will is not compatible with determinism, it does not seem to be compatible with indeterminism either. Determinism implies that, given the past and laws, there is only one possible future. Indeterminism implies the opposite: Same past and laws, different possible futures. On the face of it, indeterminism may seem more congenial to the idea of an "open" future with branching pathways in decision making. But how is it possible one might ask that

different actions or choices could arise voluntarily and intentionally for exactly the same past and (barring miraculous departures from the laws of nature) without occurring merely by luck or chance?

This question has had an hypnotic effect on those who think about free will. One imagines that if free choices were undetermined, then which occurs would be like spinning a wheel in one's mind or one must just pop out by chance or randomly. 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. Such undetermined events occurring in our brains or bodies would not seem to enhance our freedom and control over our actions, but rather diminish our freedom and control, and hence our responsibility. Arguments of these kinds and many others have led down through the centuries to familiar charges that undetermined choices, of the kind libertarians and incompatibilists demand, would be "arbitrary," "capricious," "random," "irrational," "uncontrolled," "inexplicable," "mere matters of luck or chance" and hence not really free and responsible choices at all.

No wonder libertarians about free will, who believe it is incompatible with determinism, have looked for some deus ex machina or other to solve the problem, while their opponents have cried magic or mystery. Indeterminism was required for free will, they argued, but it was not enough. It might provide causal gaps in nature. But something else must fill those gaps. Some additional form of agency or causation was needed that went beyond causation in the natural order, whether deterministic or indeterministic. Thus, in response to modern science, there were numerous historical appeals in the modern era, from Descartes to Kant and beyond, to "extra factors" such as noumenal selves, immaterial minds, transempirical power centers, non-event agent causes, uncaused causes, and the like, to account for a traditional libertarian free will. I long ago became disenchanted with all such appeals.

## **5. Indeterminism and Responsibility**

Where to go if one is to avoid such traditional libertarian strategies for explaining free will? I came to believe that one must take a whole new look at the indeterminist problem from the ground up. First, let us be clear that it is an empirical and scientific question whether any indeterminism is there in nature in ways that are appropriate for free will—in the brain, for example. It may very well not be there; and in any case, no purely philosophical theory can settle the matter. As the Epicureans said, if the atoms don't "swerve" in undetermined ways (and in the right places) there would be no room in nature for free will. I've argued for some time that the question of whether or not we have free will has an empirical dimension and

cannot be settled by a priori or philosophical reasoning alone, as philosophers have often assumed, e.g., by introspectively appealing to experiences of deliberating and choosing, or by engaging in conceptual analysis of ordinary terms like "could" or "power." That is one reason why the free will issue has been so philosophically intractable.

And yet philosophical reasoning is relevant to many aspects of the free will problem. And our present question is the philosophical one that has boggled people's minds for centuries, from the time of the Epicureans onward: 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?

The first step in addressing this question 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 would be 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 would thus be reflected in the indeterminacy of our neural processes themselves. What is experienced internally as uncertainty would correspond physically to the opening of a window of opportunity that would temporarily screen off complete determination by influences of the past.

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.

Now add a further piece to the puzzle. Just as indeterminism need not undermine rationality and voluntariness, so indeterminism in and of itself need not undermine control and responsibility. Suppose you are trying to think through a difficult problem, say a mathematical problem, and there is some indeterminacy in your neural processes complicating the task—a kind of chaotic background. It would be like trying to concentrate and solve a problem, say a mathematical problem, with background noise or distraction. Whether you are going to succeed in solving the problem is uncertain and undetermined because of the distracting neural noise. Yet, if you concentrate and solve the problem nonetheless, there is reason to say you did it and are responsible for it even though it was undetermined whether you would succeed. The indeterministic noise would have been an obstacle that you overcame by your effort.

There are numerous examples supporting this point, first suggested by J. L. Austin, Elizabeth Anscombe and others in 60s, where indeterminism functions as an obstacle to success without precluding responsibility. Consider an assassin who is trying to shoot the prime minister, but might miss because of some undetermined events in his nervous system that may lead to a wavering of his arm. If the assassin does succeed in hitting his target, despite the indeterminism, can he be held responsible? The answer is clearly yes because he intentionally and voluntarily succeeded in doing what he was trying to do—kill the prime minister. Yet his action, killing the prime minister, was undetermined. Or, here is another example: a husband, while arguing with his wife, in a fit of rage swings his arm down on her favorite glass-top table top intending to break it. Again, we suppose that some indeterminism in his outgoing neural pathways makes the momentum of his arm indeterminate so that it is genuinely undetermined whether the table will break right up to the moment when it is struck. Whether the husband breaks the table or not is undetermined and yet he is clearly responsible if he does break it. (It would be a poor excuse for him to say to his wife: "chance did it, not me." Even though there was a chance he wouldn't break it, chance didn't do it, he did.)

Now these examples—of the mathematical problem, the assassin and the husband—are not all we want, since they do not amount to genuine exercises of (self-forming) free will in

SFAs, like the businesswoman's, where the will is divided between conflicting motives. The assassin's will is not divided between conflicting motives as is the woman's. He wants to kill the prime minister, but does not also want to fail. (If he fails therefore, it will be merely by chance.) Yet these examples of the assassin, the husband and the like, do provide some clues. To go further, we have to add some further thoughts.

Imagine in cases of inner conflict characteristic of SFAs, like the businesswoman's, that the indeterministic noise which is providing an obstacle to her overcoming temptation is not coming from an external source, but is coming from her own will, since she also deeply desires to do the opposite. Imagine that two crossing (recurrent) neural networks are involved, each influencing the other, and representing her conflicting motivations. (Recurrent networks are complex networks of interconnected neurons circulating impulses in feedback loops that are generally thought to be involved in higher-level cognitive processing.<sup>v</sup>) The input of one of these neural networks consists in the woman's reasons for acting morally and stopping to help the victim; the input of the other, her ambitious motives for going on to her meeting.

The two networks are connected so that the indeterministic noise which is an obstacle to her making one of the choices is coming from her desire to make the other, and vice versa—the indeterminism thus arising from a tension-creating conflict in the will, as I said. In these circumstances, when either of the pathways reaches an activation threshold (which amounts to choice), it will be like your solving the mathematical problem by overcoming the background noise produced by the other. And just as when you solved the mathematical problem by overcoming the distracting noise, one can say you did it and are responsible for it, so one can say this as well in the present case, whichever one is chosen. The pathway through which the woman succeeds in reaching a choice threshold will have overcome the obstacle in the form of indeterministic noise generated by the other.

Note that, under such conditions, the choices either way will not be "inadvertent," "accidental," "capricious," or "merely random," (as critics of indeterminism say) because they will be willed by the agents either way when they are made, and done for reasons either way—reasons that the agents then and there endorse. But these are the conditions usually required to say something is done "on purpose," rather than accidentally, capriciously or merely by chance. Moreover, these conditions taken together, as I have argued elsewhere, rule out each of the reasons we have for saying that agents act, but do not have control over their actions (compulsion, coercion, constraint, inadvertence, accident, control by others, etc.).<sup>vi</sup>

Indeed, in these cases, agents have what I call "plural voluntary control" over the op-

tions in the following sense: they are able to bring about whichever of the options they will, when they will to do so, for the reasons they will to do so, on purpose rather than accidentally or by mistake, without being coerced or compelled in doing so or willing to do so, or otherwise controlled in doing or willing to do so by any other agents or mechanisms. I show in my 1996 book that each of these conditions can be satisfied for SFAs as conceived above even though the SFAs are undetermined.<sup>vii</sup> They can be summed up by saying, as we sometimes do, that the agents can choose either way, at will.

Note also that this account of self-forming choices amounts to a kind of "doubling" of the mathematical problem. It is as if an agent faced with such a choice is trying or making an effort to solve two cognitive problems at once, or to complete two competing (deliberative) tasks at once—in our example, to make a moral choice and to make a conflicting self-interested choice (corresponding to the two competing neural networks involved). Each task is being thwarted by the indeterminism coming from the other, so it might fail. But if it succeeds, then the agents can be held responsible because, as in the case of solving the mathematical problem, they will have succeeded in doing what they were willingly trying to do. Recall the assassin and the husband. Owing to indeterminacies in their neural pathways, the assassin might miss his target or the husband fail to break the table. But if they succeed, despite the probability of failure, they are responsible, because they will have succeeded in doing what they were trying to do.

And so it is, I suggest, with self-forming choices or SFAs, except that in the case of self-forming choices, whichever way the agents choose they will have succeeded in doing what they were trying to do because they were simultaneously trying to make both choices, and one is going to succeed. Their failure to do one thing is not a mere failure, but a voluntary succeeding in doing the other.

Does it make sense to talk about the agent's trying to do two competing things at once in this way, or to solve two cognitive problems at once? Well, much current scientific evidence points to the fact that the brain is a parallel processor; it simultaneously processes different kinds of information relevant to tasks such as perception or recognition through different neural pathways. Such a capacity, I believe, is essential to the exercise of free will. In cases of self-formation (SFAs), agents are simultaneously trying to resolve plural and competing cognitive tasks. They are, as we say, of two minds. Yet they are not two separate persons. They are not dissociated from either task. The businesswoman who wants to go back to help the victim is the same ambitious woman who wants to go to her meeting. She is torn inside by different visions of who she is and what she wants to be, as we all are from time to time. But

this is the kind of complexity needed for genuine self-formation and free will. And when she succeeds in doing one of the things she is trying to do, she will endorse that as her resolution of the conflict in her will, voluntarily and intentionally, not by accident or mistake.<sup>viii</sup>

## 6. Responsibility, Luck and Chance

Now you may find 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 charged. Such intuitions are deeply ingrained and they give rise to a host of questions and objections that naturally arise and have been made about the view just presented.

The first step in exorcising deeply ingrained intuitions about indeterminism 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 oldest and most common in debates about free will) to assume that "undetermined" means "uncaused." (Libertarian freedom was often characterized in the past as "contra-causal" freedom, which I think is misguided.)

Here is another source of misunderstanding. Since the outcome of the business-woman's effort (the choice) is undetermined up to the last minute, one 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. One cannot separate the indeterminism and the effort of will, so that first the effort occurs followed by chance or luck (or vice versa). Rather 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. Similarly, the husband breaks the table as a result of his effort, even though he might have failed because of the

indeterminacy. (That is why his excuse, "chance broke the table, not me" is so lame.)

Just as expressions like "she chose by chance" can mislead in such contexts, so can expressions like "she got lucky." Recall that, with the assassin and husband, one might say "they got lucky" in killing the prime minister and breaking the table because their actions were undetermined. Yet they were responsible. So ask yourself this question: why does the inference "he got lucky, so he was not responsible?" fail in the cases of the husband and the assassin where it does fail? The first part of an answer has to do with the point made earlier that "luck," like "chance," has question-begging implications in ordinary language that are not necessarily implications of "indeterminism." The core meaning of "he got lucky" in the assassin and husband cases, which is implied by indeterminism, I suggest, is that "he succeeded despite the probability or chance of failure"; and this core meaning does not imply lack of responsibility, if he succeeds.

The second reason why the inference "he got lucky, so he was not responsible" fails for the assassin and the husband is that what they succeeded in doing was what they were trying and wanting to do all along (kill the minister and break the table respectively). The third reason is that when they succeeded, their reaction was not "oh dear, that was a mistake, an accident—something that happened to me, not something I did." Rather they endorsed the outcomes as something they wanted all along, and did so knowingly and purposefully, not by mistake or accident.

But these conditions are satisfied in the businesswoman's case as well, either way she chooses. If she succeeds in choosing to return to help the victim (or in choosing to go on to her meeting) (i) she will have "succeeded despite the probability or chance of failure," (ii) she will have succeeded in doing what she was trying and wanting to do all along (she wanted both outcomes very much, but for different reasons, and was trying to make those reasons prevail in both cases), and (iii) when she succeeded (in choosing to return to help) her reaction was not "oh dear, that was a mistake, an accident—something that happened to me, not something I did." Rather she endorsed the outcome as her resolution of the conflict in her will. And if she had chosen to go on to her meeting she would have endorsed that outcome, recognizing it as her resolution of the conflict in her will.

Another objection often made is that we are not introspectively aware of making dual efforts and performing multiple cognitive tasks in such choice situations. But I am not claiming that agents are conscious of making dual efforts. What they are introspectively conscious of is that they are trying to decide about which of two options to choose and that either choice is a difficult one because there are resistant motives pulling them in different directions that will

have to be overcome, whichever choice is made. In such introspective conditions, I am theorizing that what is actually going on underneath is a kind of parallel distributed processing in the brain that involves separate efforts or endeavors to resolve competing cognitive tasks.

This is an example of the point made earlier that introspective evidence cannot give us the whole story about free will. Stay on the surface and things are likely to appear obscure or mysterious. What is needed is a theory about what might be going on behind the scenes when we exercise free will, not merely a description of what we immediately experience; and in this regard new scientific ideas can be a help rather than an obstacle to making sense of the subject. If parallel distributed processing takes place on the input side of the cognitive ledger (in perception), then why not consider that it also takes place on the output side (in practical reasoning, choice and action)? That is what we should suppose, I am suggesting, if we are to make sense of incompatibilist free will.

It has also been objected that indeterminism undermines the notion of agency itself by turning choices and actions into mere chance events. As noted earlier, that worry sends us scurrying around looking for extra factors, other than prior events or happenings, to tip the balance to one choice or the other. But there is an alternative way to think about the way that indeterminism might be involved in free choice, a way that avoids these familiar libertarian stratagems and requires a transformation of perspective.

The idea is not to think of the indeterminism involved in free choice as a cause acting on its own, but as an ingredient in a larger goal-directed or teleological process or activity, in which the indeterminism functions as a hindrance or obstacle to the attainment of the goal. Such is the role I have suggested for indeterminism in the efforts preceding undetermined SFAs.

We tend to reason that if an outcome (breaking a table or making a choice) depends on whether certain neurons fire or not (in the arm or in the brain), then the agent must be able to make those neurons fire or not, if the agent is to be responsible for the outcome. In other words, we think we have to crawl down to the place where the indeterminism originates (in the individual neurons) and make them go one way or the other. We think we have to become originators at the micro-level and tip the balance that chance leaves untipped, if we (and not chance) are to be responsible for the outcome. And we realize, of course, that we can't do that. But we don't have to. It's the wrong place to look. We don't have to micro-manage our individual neurons one by one to perform purposive actions and we do not have such micro-control over our neurons even when we perform ordinary actions such as swinging an arm

down on a table.

## 7. Responsibility and Control

But does not the presence of indeterminism or chance at least diminish the control persons have over their choices or actions? Is it not the case that the assassin's control over whether the prime minister is killed (his ability to realize his purposes or what he is trying to do) is lessened by the undetermined impulses in his arm—and so also for the husband and his breaking the table? And this limitation seems to be connected with another problem often cited by critics of libertarian freedom, the problem that indeterminism, wherever it occurs, seems to be a hindrance or obstacle to our realizing our purposes and hence an obstacle to our freedom rather than an enhancement of it.

There is truth 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 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). 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 no competing motives standing in the way of her choosing it and therefore no interfering indeterminism. 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. Self-formation is not a gift, but a struggle.

## 8. Liberum Arbitrium

I conclude with one final objection. 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 conclusive prior reasons for making one option and one set of reasons prevail over the other.

There is some truth to this objection 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 might be 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."

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.

Suppose we were to say to such persons: "But look, you didn't have sufficient or conclusive prior reasons for choosing as you did since you also had viable reasons for choosing the other way." They might reply. "True enough. But I did have good reasons for choosing as I did, which I'm willing to stand by and take responsibility for. If these reasons 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 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."

---

<sup>i</sup> Kane 1996: 35. The formal statement of the condition is: "(UR) An agent is ultimately responsible for some (event or state) E's occurring only if (R) the agent is personally responsible for E's occurring in a sense which entails that something the agent voluntarily (or willingly) did or omitted...either was, or causally contributed to, E's occurrence and made a difference to whether or not E occurred; and (U) for every X and Y (where X and Y represent occurrence of events and/or states) if the agent is personally responsible for X, and if Y is an arche (or sufficient ground or or explanation) for X, then the agent must also be personally responsible for Y" (1996: 35). R is the "responsibility condition" and U the "ultimacy condition" of UR. My first formulation of a condition of this sort was in Kane 1985: ch. 3. I have since dropped the phrase "for which the agent could have voluntarily done otherwise" from the statement of R in Kane 1996 because it is not needed (see Kane 2000, particularly the response to Fischer 2000); it turns out that what this phrase says follows from U for reasons discussed later in this paper. Other philosophers, such as Galen Strawson (1986) and Martha Klein (1990), have also noted the importance of an ultimacy condition for free will, though neither believes such a condition can be satisfied.

<sup>ii</sup> Aristotle 1915: 1114a13-22. Also see Richard Sorabji (1980: 234-8) for a perceptive discussion of this condition in Aristotle's writings.

<sup>iii</sup> Kane 1996: 74-8, 125 ff.

<sup>iv</sup> A proof that UR does indeed entail AP for self forming actions (SFAs) can be given. The complete version is in Kane 1996: chs. 5, 7.

<sup>v</sup> Readable and accessible introductions to the role of neural networks (including recurrent networks) in cognitive processing include P. M. Churchland, The Engine of Reason, the Seat of the Soul (MIT Press, 1996) and Manfred Spitzer, The Mind Within the Net (MIT Press, 1999). For more advanced discussion, see P. S. Churchland and T. J. Sejnowski, The Computational Brain (Cambridge MA: MIT Press, 1992).

<sup>vi</sup> We have to make further assumptions about the case to rule out some of these conditions. For example, we have to assume, no one is holding a gun to the woman's head forcing her to go back, or that she is not paralyzed, etc. But the point is that the satisfaction of these further conditions is consistent with the case of the woman as we have imagined it. If these other conditions are satisfied, as they can be, and the businesswoman's case is in other respects as I have described it, we have an SFA. I offer the complete argument for this in The Significance of Free Will, chapter 8, among other works listed in note 3.

<sup>vii</sup> I show in greater detail that each of these conditions can be satisfied by SFAs in The Significance of Free Will, chapter 8-10.

<sup>viii</sup> Another related objection that is commonly made at this point is that it is irrational to make efforts to do incompatible things. I concede that in most ordinary situations it is. But I argue that there are special circumstances that can arise in the deliberative lives of rational agents in which

---

it is not irrational to make competing efforts: These include circumstances in which (i) we are deliberating between competing options; (ii) we intend to choose one or the other, but cannot choose both; (iii) we have powerful motives for wanting to choose each of the options for different and incommensurable reasons; (iv) there is a consequent resistance in our will to either choice, so that (v) if either choice is to have a chance of being made, effort will have to be made to overcome the temptation to make the other choice; and most importantly, (vi) we want to give each choice a fighting chance of being made because the motives for each choice are important to us. The motives for each choice define in part what sort of person we are; and we would taking them lightly if we did not make an effort in their behalf. These conditions are the conditions of SFAs.