

---

---

# THE JOURNAL OF PHILOSOPHY

VOLUME XCVI, NO. 5, MAY 1999

---

---

## RESPONSIBILITY, LUCK, AND CHANCE: REFLECTIONS ON FREE WILL AND INDETERMINISM\*

Ludwig Wittgenstein<sup>1</sup> once said that “to solve the problems of philosophers, you have to think even more crazily than they do” (*ibid.*, p. 75). This task (which became even more difficult after Wittgenstein than it was before him) is certainly required for the venerable problem of free will and determinism.

### I. THE LUCK PRINCIPLE

Consider the following principle:

(LP) If an action is *undetermined* at a time *t*, then its happening rather than not happening at *t* would be a matter of *chance* or *luck*, and so it could not be a *free* and *responsible* action.

This principle (which we may call the *luck principle*, or simply LP) is false, as I shall explain shortly. Yet it seems true. LP and a related principle to be considered later in this paper are fueled by many of those “intuition pumps,” in Daniel Dennett’s<sup>2</sup> apt expression, which

\* This paper was prompted by a recent objection made in various forms against my view and other incompatibilist views of freedom and responsibility by Galen Strawson, Alfred Mele, Bernard Berofsky, Bruce Waller, Richard Double, Mark Bernstein, and Ishtiyaque Haji. (See footnote 10 for references.) The paper has benefitted from interchanges with the above persons and with participants at a conference on my work on free will at the University of Arkansas in September, 1997: Gary Watson, Barry Loewer, Timothy O’Connor, Randolph Clarke, Christopher Hill, and Thomas Senor. It has also benefitted from interchanges in conferences or in correspondence with John Martin Fischer, William Rowe, Nicholas Nathan, David Hodgson, Saul Smilansky, Kevin Magill, Peter van Inwagen, Derk Pereboom, Laura Ekstrom, Hugh McCann, and Ilya Prigogine. I am especially grateful to Mele and Strawson for pursuing me assiduously on these issues since the publication of my latest work, and for perceptive comments on the penultimate draft by Mele, Berofsky, and George Graham.

<sup>1</sup> *Culture and Value* (New York: Blackwell, 1980).

<sup>2</sup> *Elbow Room* (Cambridge: MIT, 1984), chapter 1 and pp. 32-34, 64-65, 119-20, 169-70.

support common intuitions about freedom and responsibility. LP and related principles lie behind the widespread belief that indeterminism, so far from being required for free will and responsibility, would actually undermine free will and responsibility. Dennett does not dwell on the intuition pumps of this sort, as I shall do in this paper. As a compatibilist, he is more interested in criticizing intuition pumps that lead people to think (mistakenly, on his view) that freedom and responsibility are not compatible with determinism, whereas intuition pumps that support LP lead people to think freedom and responsibility are not compatible with *indeterminism*. Yet intuition pumps of the latter kind are every bit as pervasive and influential in free-will debates as those Dennett dwells upon; and they are as much in need of deconstruction, since they play a significant role in leading people to believe that freedom and responsibility must be compatible with determinism.

I think the modern route to compatibilism—which is the reigning view among contemporary philosophers—usually goes through principles like LP at some point or other. In my experience, most ordinary persons start out as natural incompatibilists. They believe there is some kind of conflict between freedom and determinism; and the idea that freedom and responsibility might be compatible with determinism looks to them at first like a “quagmire of evasion” (William James) or “a wretched subterfuge” (Immanuel Kant). Ordinary persons have to be talked out of this natural incompatibilism by the clever arguments of philosophers—who, in the manner of their mentor, Socrates, are only too happy to oblige. To weaken natural incompatibilist instincts, philosophers first argue that what we mean by freedom in everyday life is the power or ability to do whatever we choose or desire to do—in short, an absence of coercion, compulsion, oppression, and other impediments or constraints upon our behavior. They then point out that we can be free in these everyday senses to do what we choose or desire, even if our choices and desires are determined by causes that lie in our past.

But this line of argument does not usually dispose of incompatibilist intuitions by itself. Ordinary persons might grant that many everyday freedoms are compatible with determinism and still wonder if there is not also some deeper freedom—the freedom to have an *ultimate* say in what we choose or desire to do in the first place—that is incompatible with determinism. (I have argued elsewhere<sup>3</sup> that this deeper freedom is what was traditionally meant by “free will.”) So the philosophers must add a second step to their case—an argu-

<sup>3</sup> *The Significance of Free Will* (New York: Oxford, 1996), pp. 10-14, 33-37.

ment to the effect that any allegedly deeper freedom (of the will) that is not compatible with determinism is no intelligible freedom at all. And with this step, principles like LP come into the picture. For any freedom not compatible with determinism would require indeterminism; and what is undetermined, it seems, would happen by chance or luck and could not be a free and responsible action. This kind of argument is the one that usually puts the final nail in the coffin of incompatibilist instincts.

When philosophy professors go through this two-stage argument in the modern classroom, they are replicating the standard case against traditional (incompatibilist or libertarian) free will which is one of the defining characteristics of modernity. The goal is to consign incompatibilist freedom to the dustbin of history with other beliefs that a modern scientific age is encouraged to outgrow. Students and ordinary persons subjected to this argument may have an uneasy feeling they are being had by the clever arguments of philosophers. But, also seeing no obvious response, except an appeal to mystery, many of them become compatibilists.

## II. INDETERMINISM, THE BOGEYMAN

The second stage of this two-stage argument in support of compatibilism will concern me here, the one that goes through LP and related principles in the attempt to show that indeterminism would not enhance, but in fact would undermine, freedom and responsibility. What is at stake here is not merely the clever arguments of philosophers; for it happens that the case for principles like LP is a powerful one. It *is* difficult to see how indeterminism and chance can be reconciled with freedom and responsibility. Philosophers have tried to bring this out in a number of ways which will be addressed here. We may think of these as the varied intuition pumps that support LP and principles like it.

(1) We are often asked to consider, for example, that whatever is undetermined or happens by chance is not under the *control* of anything, and so is not under the control of the agent. But an action that is not under the control of the agent could not be a free and responsible action. (Here it is evident that the notion of control is involved in the case for LP: indeterminism and chance imply lack of control to a degree that implies lack of freedom and responsibility.)

(2) Another line of argument often heard is this: suppose a choice occurred as the result of an undetermined event (say, a quantum jump) in one's brain. Would that be a free choice? Being undetermined, it would appear to be more of a fluke or accident than a free and responsible action. Some twentieth-century scientists and

philosophers have suggested that free will might be rescued by supposing that undetermined quantum events in the brain could be amplified to have large-scale effects on choice or action.<sup>4</sup> Unfortunately, this modern version of the ancient Epicurean “swerve” of the atoms seems to be subject to the same criticisms as its ancient counterpart. It seems that undetermined events in the brain or body, whether amplified or not, would occur spontaneously and would be more of a nuisance—or perhaps a curse, like epilepsy—than an enhancement of freedom and responsibility.

(3) Nor would it help to suppose that the indeterminism or chance came *between* our choices (or intentions) and our actions. Imagine that you are intending to make a delicate cut in a fine piece of cloth, but because of an undetermined twitching in your arm, you make the wrong cut. Here, indeterminism is no enhancement of your freedom, but a *hindrance* or *obstacle* to your carrying out your purposes as intended. Critics of libertarian freedom<sup>5</sup> have often contended that this is what indeterminism would always be—a hindrance or impediment to one’s freedom. It would get in the way, *diminishing* control, and hence responsibility, instead of enhancing them.

(4) Even more absurd consequences follow if we suppose that indeterminism or chance is involved in the initiation of overt actions. Arthur Schopenhauer<sup>6</sup> imagined the case of a man who suddenly found his legs start to move *by chance*, carrying him across the room against his wishes. Such caricatures are popular among critics of indeterminist freedom for obvious reasons: undetermined or chance-initiated overt actions would represent the opposite of controlled and responsible actions.

(5) Going a little deeper, one may also note that, if a choice or action is undetermined, it might occur otherwise *given exactly the same past and laws of nature* up to the moment when it does occur. This means that, if Jane is deliberating about whether to vacation in Hawaii or Colorado, and gradually comes to favor and choose Hawaii, she might have chosen otherwise (chosen Colorado), given *exactly the same deliberation* up to the moment of choice that in fact led

<sup>4</sup> For example, physicist A. H. Compton, *The Freedom of Man* (New Haven: Yale, 1935) and neurophysiologist John Eccles, *Facing Reality* (New York: Springer, 1970).

<sup>5</sup> See, for example, Galen Strawson, who argues that, even if free will should be incompatible with determinism, indeterminism would be “no help” in enhancing either freedom or responsibility—“The Unhelpfulness of Indeterminism,” *Philosophy and Phenomenological Research* (forthcoming).

<sup>6</sup> *Essay on the Freedom of the Will* (Indianapolis: Bobbs-Merrill, 1960), p. 47.

her to favor and choose Hawaii (exactly the same thoughts, reasonings, beliefs, desires, dispositions, and other characteristics—not a sliver of difference). It is difficult to make sense of this. The choice of Colorado in such circumstances would seem irrational and inexplicable, capricious and arbitrary.<sup>7</sup> If it came about by virtue of undetermined events in Jane's brain, this would not be an occasion for rejoicing in her freedom, but for consulting a neurologist about the waywardness of her neural processes.

(6) At this point, some defenders of incompatibilist freedom appeal to Gottfried Leibniz's<sup>8</sup> celebrated dictum that prior reasons or motives need not determine choice or action, they may merely "incline without necessitating"—that is, they may incline the agent toward one option without determining the choice of that option. This may indeed happen. But it will not solve the present problem; for it is precisely *because* Jane's prior reasons and motives (beliefs, desires, and the like) incline her toward the choice of Hawaii that choosing Colorado by chance at the end of exactly the same deliberation would be irrational and inexplicable. Similarly, if her reasons had inclined her toward Colorado, then choosing Hawaii by chance at the end of the same deliberation would have been irrational and inexplicable. And if prior reasons or motives had not inclined her either way (the celebrated medieval "liberty of indifference") and the choice was a matter of chance, then the choosing of one rather than the other would have been all the more a matter of luck and out of her control. (One can see why libertarian freedom has often been ridiculed as a mere "liberty of indifference.")

(7) Indeed, critics of indeterminist freedom have often argued that indeterminist free choices must always amount to *random* choices of this sort and hence the outcomes would be matters of mere luck or chance—like spinning a wheel to select among a set of alternatives. Perhaps there is a role for such random choices in our lives when we are genuinely indifferent to outcomes.<sup>9</sup> But to suppose that *all* of our free and responsible choices—including momentous ones, like whether to act heroically or treacherously—had to be by random selection in this way has been regarded by many philosophers as a *reductio ad absurdum* of the view that free will and responsibility require indeterminism.

<sup>7</sup> This dilemma for incompatibilist accounts of freedom is nicely described by Thomas Nagel, *The View from Nowhere* (New York: Oxford, 1986), chapter 7.

<sup>8</sup> *Selections* (New York: Scribner's, 1951), p. 435.

<sup>9</sup> Stephen M. Cahn makes a persuasive case for there being such a role—"Random Choices," *Philosophy and Phenomenological Research*, xxxvii (1977): 549-51.

(8) Consider one final argument which cuts more deeply than the others and to which I shall devote considerable attention. This paper was in fact prompted by new versions of this argument advanced in recent years against my incompatibilist account of free will by Galen Strawson, Alfred Mele, Bernard Berofsky, Bruce Waller, Richard Double, Mark Bernstein, and Ishtiyaque Haji<sup>10</sup>—though the argument is meant to apply generally to any view requiring that free actions be undetermined up to the moment when they occur.

Suppose two agents had exactly the same pasts (as indeterminism requires) up to the point where they were faced with a choice between distorting the truth for selfish gain or telling the truth at great personal cost. One agent lies and the other tells the truth. As Waller puts it, if the pasts of these two agents “are really identical” in every way up to the moment of choice, “and the difference in their acts results from chance,” would there “be any grounds for distinguishing between [them], for saying that one deserves censure for a selfish decision and the other deserves praise” (*op. cit.*, p. 151)? Mele poses the problem in terms of a single agent in different possible worlds. Suppose in the actual world, John fails to resist the temptation to do what he thinks he ought to do, arrive at a meeting on time. If he could have done otherwise given the same past, then his counterpart, John\* in a nearby possible world, which is the same as the actual world up to the moment of choice, resists the temptation and arrives on time. Mele then argues that, “if there is nothing about the agents’ powers, capacities, states of mind, moral character and the like that explains this difference in outcome,...the difference is just a matter of luck” (*op. cit.*, pp. 582-83). It would seem that John\* got lucky in his attempt to overcome temptation, whereas John did not. Would it be just to reward the one and punish the other for what appears to be ultimately the luck of the draw?

<sup>10</sup> Strawson, “The Impossibility of Moral Responsibility,” *Philosophical Studies*, LXXV (1994): 5-24, and “The Unhelpfulness of Indeterminism,” *op. cit.*; Mele, Review of my *The Significance of Free Will*, this JOURNAL, XCV, 11 (November 1998): 581-84, and “Luck and the Significance of Free Will,” *Philosophical Explorations* (forthcoming); Berofsky, “Ultimate Responsibility in a Deterministic World,” *Philosophy and Phenomenological Research* (forthcoming); Waller, “Free Will Gone Out of Control,” *Behaviorism*, XVI (1988): 149-67; Double, *The Non-reality of Free Will* (New York: Oxford, 1991), p. 140; Bernstein, “Kanean Libertarianism,” *Southwest Philosophy Review*, XI (1995): 151-57; Haji, “Indeterminism and Frankfurt-type Examples,” *Philosophical Explorations* (forthcoming). Different, but related, concerns about indeterminism and agency are aired by Timothy O’Connor, “Indeterminism and Free Agency: Three Recent Views,” *Philosophy and Phenomenological Research*, LIII (1993): 499-526; and Randolph Clarke, “Free Choice, Effort and Wanting More,” *Philosophical Explorations* (forthcoming).

Considerations such as (1)-(8) lie behind familiar and varied charges that undetermined choices or actions would be “arbitrary,” “capricious,” “random,” “uncontrolled,” “irrational,” “inexplicable,” or “matters of luck or chance,” and hence not free and responsible actions. These are the charges which principles like LP are meant to express. Responses to them in the history of philosophy have been many; but none to my mind has been entirely convincing. The charges have often led libertarians—those who believe in an incompatibilist free will—to posit “extra factors” in the form of unusual species of agency or causation (such as noumenal selves, immaterial egos, or nonoccurrent agent causes) to account for what would otherwise be arbitrary, uncontrolled, inexplicable, or mere luck or chance. I do not propose to appeal to any such extra factors in defense of libertarian freedom. Such appeals introduce additional problems of their own without, in my view, directly confronting the deep problems about indeterminism, chance, and luck to which considerations (1)-(8) are pointing. To confront these deep problems directly, I believe one has to rethink issues about indeterminism and responsibility from the ground up, without relying on appeals to extracausal factors—a task to which I now turn.

### III. INDETERMINISM AND RESPONSIBILITY

First, one must question the intuitive connection in people’s minds between “indeterminism’s being involved in something’s happening” and “its happening merely as a matter of chance or luck.” ‘Chance’ and ‘luck’ are terms of ordinary language which carry the connotation of “its being out of my control” (as in (1) and (4) and above). So using them already begs certain questions, whereas ‘indeterminism’ is a technical term that merely precludes *deterministic* causation (though not causation altogether). Second, one must emphasize that indeterminism does not have to be involved in all free and responsible acts, even for incompatibilists or libertarians.<sup>11</sup> Frequently, we act from a will already formed; and it may well be that our actions are determined in such cases by our

<sup>11</sup> I defend this point at length in *Free Will and Values* (Albany: SUNY, 1985), chapters 4 and 5. It is also defended by Peter van Inwagen, “When Is the Will Free?” in J. Tomberlin, ed., *Philosophical Perspectives*, Volume 3 (Atascadero, CA: Ridgeview, 1989), pp. 399-422. John Martin Fischer has described the view that van Inwagen and I defend as “restricted libertarianism,” and has criticized it in “When the Will Is Free,” in Tomberlin, ed., *Philosophical Perspectives*, Volume 6 (Atascadero: Ridgeview, 1992), pp. 423-51. Another critic is Hugh McCann, “On When the Will Is Free,” in G. Holmstrom-Hintikka and R. Tuomela, eds., *Contemporary Action Theory*, Volume 1 (Dordrecht: Kluwer, 1997), pp. 219-32. van Inwagen responds to Fischer in “When Is the Will Not Free?” *Philosophical Studies*, LXXV (1994): 95-114; and I respond in *The Significance of Free Will*, pp. 32-43.

then existing characters and motives. On such occasions, to do otherwise by chance *would* be a fluke or accident, irrational and inexplicable, as critics of indeterminist freedom contend (in (3) and (4) above).

Incompatibilists about free will should not deny this. What they should rather say is that when we act from a will already formed (as we frequently do), it is “our own free will” by virtue of the fact that we formed it (at least in part) by earlier choices or actions which were not determined and for which we could have done otherwise voluntarily, not merely as a fluke or accident. I call these earlier undetermined actions *self-forming actions* or SFAs.<sup>12</sup> Undetermined SFAs are a subset of all of the actions done of our own free wills (many of which may be determined by our earlier formed character and motives). But if there were no such undetermined SFAs in our lifetimes, there would have been nothing we could have ever voluntarily done to make ourselves different than we are—a condition that I think is inconsistent with our having the kind of responsibility for being what we are which genuine free will requires.

Now, let us look more closely at these undetermined SFAs. As I see it, they occur at times in 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 self-interest, or between present desires and long-term goals, or we are faced with difficult tasks 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. In the light of this picture, I suggest the following incompatibilist account of SFAs.<sup>13</sup> There is a tension and uncertainty in our minds at such times of inner conflict which are 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. As a result, the uncertainty and inner tension we feel at such soul-searching moments of self-formation is reflected in the indeterminacy of our neural processes themselves. What is experienced phenomenologically as uncertainty corresponds physically to the opening of a win-

<sup>12</sup> See *The Significance of Free Will*, pp. 74-78. SFAs are also sometimes called “self-forming willings” or SFWs in that work (pp. 125ff.).

<sup>13</sup> This, in broad outline, is the account developed in my *The Significance of Free Will*, chapters 8-10. In later sections below, I make important additions to it in response to criticisms.



dow of opportunity that temporarily screens off complete determination by the past. (By contrast, when we act from predominant motives or settled dispositions, the uncertainty or indeterminism is muted. If it were involved then, it *would* be a mere nuisance or fluke, capricious or arbitrary, as critics contend (in (2), (5) and (6) above).)

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. If we overcome temptation, it will be the result of our effort; and if we fail, it will be because we did not *allow* our effort to succeed. And this is owing to the fact that, while we wanted to overcome temptation, we also wanted to fail, for quite different and incommensurable reasons. When we 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*.

Return now to concerns about indeterminism and responsibility in the light of this picture. Consider a businesswoman who faces a conflict in her will of the kind typically involved in such SFAs. She is on the way to a meeting important to her career when she observes an assault in an alley. An inner struggle ensues between her moral conscience, to stop and call for help, and her career ambitions, which tell her she cannot miss this meeting—a struggle she eventually resolves by turning back to help the victim. Now suppose this woman visits some future neuroscientists the next day and they tell her a story about what was going on in her brain at the time she chose, not unlike the story just told. Prior to choice, there was some indeterminacy in her neural processes stirred up by the conflict in her will. The indeterminism made it uncertain (and undetermined) whether she would go back to help or press onward.

Suppose further that two recurrent and connected neural networks are involved in the neuroscientists' story. Such networks circulate impulses and information in feedback loops and generally play a role in complex cognitive processing in the brain of the kind that one would expect to be involved in human deliberation. Moreover, recurrent networks are nonlinear, thus allowing (as some recent research suggests) for the possibility of chaotic activity, which would contribute to the plasticity and flexibility human brains display in

creative problem solving (of which practical deliberation is an example).<sup>14</sup> The input of one of these recurrent networks consists of the woman's moral motives, and its output the choice to go back; the input of the other, her career ambitions, and its output, the choice to go on to her meeting. The two networks are connected, so that the indeterminism that made it uncertain that she would do the moral thing was coming from her desire to do the opposite, and vice versa—the indeterminism thus arising, as we said, from a conflict in the will. When her effort to overcome self-interested desires succeeded, this corresponded to one of the neural pathways reaching an activation threshold, overcoming the indeterminism generated by the other.

To this picture, one might now pose the following objection: if it really was undetermined which choice the woman would make (in neural terms, which network would activate) right up to the moment when she chose, it seems that it would be a matter of luck or chance that one choice was made rather than the other, and so she could not be held responsible for the outcome. (Note that this is an expression of LP.) The first step in response is to recall a point made earlier: we must be wary of moving too hastily from 'indeterminism is involved in something's happening' to 'its happening merely as a matter of chance or luck'. 'Luck' and 'chance' have meanings in ordinary language that mere indeterminism may not have. The second step is to note that indeterminism of itself does not necessarily undermine control and responsibility.<sup>15</sup> 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 with background noise or distraction. Whether you are going to succeed in solving the mathematical problem is uncertain and undetermined because of the distracting neural noise. Yet if you concentrate and solve the problem nonetheless, I think we can say that you did it and are responsible for doing

<sup>14</sup> See P. Huberman and G. Hogg, "Phase Transitions in Artificial Intelligence Systems," *Artificial Intelligence*, xxxiii (1987): 155-72; C. Skarda and W. Freeman, "How Brains Make Chaos in Order to Make Sense of the World," *Behavior and Brain Sciences*, x (1987): 161-95; A. Babloyantz and A. Destexhe, "Strange Attractors in the Human Cortex," in L. Rensing, ed., *Temporal Disorder in Human Oscillatory Systems* (New York: Springer, 1985), pp. 132-43.

<sup>15</sup> Important recent defenses of the claim that indeterminism does not necessarily undermine control and responsibility include Clarke, "Indeterminism and Control," *American Philosophical Quarterly*, xxxii (1995): 125-38; Carl Ginet, *On Action* (New York: Cambridge, 1990), chapter 6; O'Connor; and Laura Ekstrom, *Free Will* (Boulder: Westview, forthcoming).

it even though it was undetermined whether you would succeed. The indeterministic noise would have been an obstacle to your solving the problem which you nevertheless overcame by your effort.

There are numerous other examples in the philosophical literature of this kind, where indeterminism functions as an obstacle to success without precluding responsibility. Consider an assassin who is trying to kill the prime minister, but might miss because of some undetermined events in his nervous system which might lead to a jerking or wavering of this arm. If he does hit his target, can he be held responsible? The answer (as J. L. Austin and Philippa Foot<sup>16</sup> successfully argued decades ago) is “yes,” because he intentionally and voluntarily succeeded in doing what he was *trying* to do—kill the prime minister. Yet his killing the prime minister was undetermined. We might even say in a sense that he got lucky in killing the prime minister, when he could have failed. But it does not follow, if he succeeds, that killing the prime minister was not his action, not something he did; nor does it follow, as LP would require, that he was not responsible for killing the prime minister. Indeed, if anything is clear, it is that he both killed the prime minister and was responsible for doing so.

Or consider a husband who, while arguing with his wife, swings his arm down in anger on her favorite glass table top, intending to break it. Again we suppose that some indeterminism in the husband’s efferent neural pathways makes the momentum of his arm indeterminate, so it is undetermined if the table will break right up to the moment when it is struck. Whether the husband breaks the table or not is undetermined. Yet it does not follow, if he succeeds, that breaking the table was not something he did; nor again does it follow, as LP would require, that he was not responsible for breaking it.<sup>17</sup> The inference sanctioned by LP from ‘it was undetermined’ to ‘he was not responsible’, is not valid. The above cases are counterexamples to it; and there are many more.

#### IV. POSSIBLE WORLDS AND LP\*

But one may grant this and still object that counterexamples to LP of these kinds do not amount to genuine exercises of free will involving

<sup>16</sup> Austin, “Ifs and Cans,” in his *Philosophical Papers* (New York: Oxford, 1961), pp. 153-80; Foot, “Free Will as Involving Determinism,” in Berofsky, ed., *Free Will and Determinism* (New York: Harper and Row, 1966), pp. 95-108.

<sup>17</sup> We must, of course, assume in both these examples that other (compatibilist) conditions for responsibility are in place—for example, that, despite his anger, the husband was not acting compulsively and would have controlled himself, if he had wished; that he knew what he was doing and was doing it intentionally to anger his wife, and so on (and similarly for the assassin). But the point is that nothing in the facts of either case preclude these assumptions from also being satisfied.

SFAs, such as the businesswoman's, where there is conflict in the wills of the agents and they are supposed to choose freely and responsibly *whichever* way they choose. If the assassin and husband succeed in doing what they are trying to do (kill the prime minister, break the table) they will do it *voluntarily* (in accordance with their wills) and *intentionally* (knowingly and purposely). But if they *fail* because of the indeterminism, they will not fail voluntarily and intentionally, but "by mistake" or "accident," or merely "by chance." Thus, their "power" to do *otherwise* (if we should even call it a power) is not the usual power we associate with freedom of choice or action in self-formation, where the agents should be able to choose or act either way voluntarily or intentionally. The power to do otherwise of the assassin and the husband is more like Jane's "power" in (5) and (6) of section II, to choose to vacation in Colorado by a fluke or accident, after a long deliberation in which she had come to favor Hawaii.

As a consequence, while LP may fail for cases like those of the assassin, husband, and mathematical problem solver, another luck principle similar to LP might still be applicable to genuine exercises of free will involving SFAs, like the businesswoman's: if it is undetermined at *t* whether an agent *voluntarily* and *intentionally* does *A* at *t* or *voluntarily* and *intentionally* does otherwise, then the agent's doing one of these rather than the other at *t* would be a matter of *luck* or *chance*, and so could not be a free and responsible action. This principle—let us call it LP\*—is fueled by the same intuitions that fuel LP. Indeed, it is a special case of LP, but one that is more difficult to deal with because it is not subject to counterexamples like those of the husband and the assassin; and it seems to be applicable to SFAs, like the businesswoman's, where failure is not merely a matter of mistake or accident.

To explore further the difficulties posed by LP\*, let us look at the final and, I think, most powerful of the intuition pumps in support of LP-type principles mentioned in section II, namely, consideration (8). This was the argument of Strawson, Mele, Berofsky, Waller, Double, Bernstein, and Haji about two agents, or one agent in different possible worlds, with the same pasts.

Consider the version of this argument by Mele, which appeared in this JOURNAL and is a particularly revealing and challenging version of it. In the actual world, an agent John succumbs to the temptation to arrive late to a meeting, whereas his counterpart, John\*, in a nearby possible world, whose physical and psychological history is the same as John's up to the moment of choice (as indeterminism requires), resists this temptation. Similarly, we can imagine a coun-

terpart to the businesswoman, businesswoman\*, in a nearby possible world who goes to her meeting instead of stopping to aid the assault victim, given the same past. But then, Mele argues, “if there is nothing about [these] agents’ powers, capacities, states of mind, moral character and the like that explains this difference in outcome,” since they are the same up to the moment of choice in the two possible worlds, “then the difference is just a matter of luck” (*op. cit.*, p. 583).<sup>18</sup> It would seem that John\* got lucky in his attempt to overcome temptation, whereas John did not; and similarly, the businesswoman got lucky in her attempt to overcome temptation, while businesswoman\* did not.

Let us first consider a general form of this argument that would support LP.

- (a) In the actual world, person *P* (for example, John, the businesswoman) does *A* at *t*.

On the assumption that the act is undetermined at *t*, we may imagine that:

- (b) In a nearby-possible world which is the same as the actual world up to *t*, *P*\* (*P*’s counterpart with the same past) does otherwise (does *B*) at *t*.
- (c) But then (since their pasts are the same), there is nothing about the agents’ powers, capacities, states of mind, characters, dispositions, motives, and so on prior to *t* which explains the difference in choices in the two possible worlds.
- (d) It is therefore a matter of luck or chance that *P* does *A* and *P*\* does *B* at *t*.
- (e) *P* is therefore not responsible (praiseworthy or blameworthy, as the case may be) for *A* at *t* (and presumably *P*\* is also not responsible for *B*).

Call this the *luck argument*. The key assumption is the assumption of indeterminism, which leads to step (b). The remaining steps are meant to follow from (b), given (a).

Despite the fact that this argument looks like Mele’s and has an initial plausibility, it is not his argument—and it is a good thing it is not. For the argument from (a)-(e) is invalid as it stands—for the same reasons that LP was invalid. Consider the husband and hus-

<sup>18</sup> I have elsewhere denied that the pasts of the agents can be exactly the same, since, with indeterminist efforts, there is no exact sameness or difference (*The Significance of Free Will*, pp. 171-74). Mele’s argument is designed to work, however, whether this denial of exact sameness is assumed or not. So I do not make an issue of it here.

band\* (his counterpart in a nearby world who fails to break the wife's table). If the outcome is undetermined, husband and husband\* also have "the same powers, capacities, states of mind, characters, dispositions, motives, and so on" up to the moment of breaking or not breaking the table, as the argument requires; and it is a matter of luck or chance that the table breaks in one world and not the other. But for all that, it does not follow, as (e) requires, that the husband is not responsible for breaking the table. The husband would have quite a task persuading his wife that he was not responsible for breaking the table on the grounds that it was a matter of luck or chance that it broke. ("Luck or chance did it, not me" is an implausible excuse.)

But, of course, as we noted, husband\* is not also responsible for *failing* to break the table, since he does not fail to break it voluntarily or intentionally. He is responsible only for the attempt, when he fails. Similarly, assassin\* would be responsible for the attempted murder of the prime minister, when he missed. What has to be explicitly added to the argument (a)-(e) to avoid counterexamples like these is the LP\* requirement that *both P* and *P\** *voluntarily* and *intentionally* do *A* and *B* respectively in their respective worlds. Specifically, we must add to premise (a) that *P* voluntarily and intentionally does *A* at *t* and to (b), that *P\** voluntarily and intentionally does *B* at *t*, and then make the corresponding additions to (d) and (e). This will yield what we might call the LP\* version of the luck argument rather than the LP version. And the stronger LP\* version is clearly the one Mele intends, since John's choice in his example is supposed to be an SFA, like the businesswoman's choice in my example, where the agents can go either way voluntarily and intentionally. Moreover, this version of the argument—like LP\* itself—is immune to counterexamples like those of the husband and the assassin.

#### V. PARALLEL PROCESSING

Nonetheless, despite immunity from these counterexamples, I think the LP\* version of the luck argument, and LP\* itself, also fail. But it is far less easy to show why. To do so, we have to take a closer look at SFAs and push the argument beyond where it has come thus far. Let it be granted that the businesswoman's case and other SFAs like John's are not like the examples of the husband and the assassin. The wills of the husband and assassin are already "set" on doing what they intend, whereas the wills of agents in SFAs, like the businesswoman and John, are not already settled or "formed" until they choose (hence the designation "self-forming actions").<sup>19</sup>

<sup>19</sup> See *The Significance of Free Will*, pp. 112-14.

Thus, to get from examples like those of the husband and assassin to genuine SFAs, I think we must do two things. First, we must put the indeterminacy involved in the efferent neural pathways of the husband and assassin into the central neural processes of the businesswoman and other agents, like John, who are making efforts of will to overcome moral, prudential, and other temptations. This move has already been made in earlier sections. But to respond to LP\* versions of the luck argument, like Mele's, I believe this move must also be combined with another—a kind of “doubling” of the example given earlier of solving the mathematical problem in the presence of background indeterministic noise.<sup>20</sup>

Imagine that the businesswoman is *trying* or making an effort to solve *two* cognitive problems at once, or to complete two competing (deliberative) tasks at once—to make a moral choice and to make a choice for her ambitions (corresponding to the two competing neural networks involved in the earlier description). With respect to each task, as with the mathematical problem, she is being thwarted in her attempt to do what she is trying to do by indeterminism. But in her case, the indeterminism does not have a mere external source; it is coming from her own will, from her desire to do the opposite. Recall that the two crossing neural networks involved are connected, so that the indeterminism which is making it uncertain that she will do the moral thing is coming from her desire to do the opposite, and vice versa. She may therefore fail to do what she is trying to do, just like the assassin, the husband, and the person trying to solve the mathematical problem. But I argue that, if she nevertheless *succeeds*, then she can be held responsible because, like them, she will have succeeded in doing *what she was trying to do*. And the interesting thing is that this will be true of her, *whichever choice is made*, because she was trying to make both choices and one is going to succeed.

Does it make sense to talk about agents trying to do two competing things at once in this way? Well, we know the brain is a parallel processor and that capacity, I believe, is essential for the exercise of free will. In cases of self-formation, agents are simultaneously trying to resolve plural and competing cognitive tasks. They are, as we say,

<sup>20</sup> This further “doubling” move is consistent with the theory put forward in *The Significance of Free Will*, and presupposes much of that theory, but is not made in that work. It is a further development especially provoked by Mele's argument discussed here as well as by criticisms of other persons since the book's publication, such as Strawson, Berofsky, Nicholas Nathan, Gary Watson, Clarke, O'Connor, Double, and Haji.

of two minds. But they are not therefore two separate persons. They are not disassociated from either task.<sup>21</sup> The businesswoman who wants to go back and help the assault victim is the same ambitious woman who wants to go on to her meeting and close the sale. She is a complex creature, like most of us who are often torn inside; but hers is the kind of complexity needed for 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, as LP\* requires. She will not disassociate from either outcome, as did Jane (in (5) of section II), who wondered what “happened to” her when she chose Colorado, or like the husband and assassin who did not also want to fail.<sup>22</sup>

But one may still object that the businesswoman makes one choice rather than the other *by chance*, since it was undetermined right up to the last moment which choice she would make. If this is so, we may have the picture of her first making an effort to overcome temptation (to go on to her meeting) and do the moral thing, and then at the last minute “chance takes over” and decides the issue for her. But this is the wrong picture. On the view just described, you cannot separate the indeterminism from the effort to overcome temptation in such a way 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 woman’s effort of will has this property of being indeterminate does not make it any less her *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

<sup>21</sup> I account for this elsewhere in terms of the notion of a “self-network” (*The Significance of Free Will*, pp. 137-42), a more comprehensive network of neural connections representing the general motivational system in terms of which agents define themselves as agents and practical reasoners. For further discussion of such a notion, see Owen Flanagan, *Consciousness Reconsidered* (Cambridge: MIT, 1992), pp. 207ff.

<sup>22</sup> In response to my claim (*The Significance of Free Will*, p. 215) that “free willers [who engage in SFAs] are always trying to be better than they are by their own lights,” by trying to overcome temptations of various sorts, Strawson asks: but “can’t they also try to be worse than they are?”—“The Unhelpfulness of Indeterminism.” He is right, of course; they can. I should have added what I am saying here, that free willers can and do *also* try to be as bad or worse than they are by resisting efforts to be better. Strange creatures indeed.



chooses *as a result of* the effort, even though she might have failed because of the indeterminism.

And just as expressions like 'She chose *by chance*' can mislead us in these contexts, so can expressions like 'She got lucky'. Ask yourself this question: Why does the inference 'He got lucky, *so he was not responsible*' fail when it does fail, as in the cases of the husband and the assassin? The first part of an answer goes back to the claim that 'luck', like 'chance', has question-begging implications in ordinary language which are not necessarily implications of "indeterminism" (which implies only the absence of deterministic causation). The core meaning of 'He got lucky', 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.

If 'He got lucky' had further meanings in these contexts often associated with 'luck' and 'chance' in ordinary usage (for example, the outcome was not his doing, or occurred by *mere chance*, or he was not responsible for it), the inference would not fail for the husband and assassin, as it clearly does. But the point is that these further meanings of 'luck' and 'chance' do not follow *from the mere presence of indeterminism*. Second, the inference 'He got lucky, so he was not responsible' fails because *what* the assassin and husband succeeded in doing was what they were trying and wanting to do all along. Third, *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 were trying and wanting to do all along, that is to say, 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 something she was trying and wanting to do all along; she recognized it as her resolution of the conflict in her will. And if she had chosen to go on to her meeting she would have en-

dorsed that outcome, recognizing it as her resolution of the conflict in her will.

#### VI. THE LUCK ARGUMENT REVISITED

With this in mind, let us return to the LP\* version of the argument from (a)-(e). I said that Mele clearly intends this stronger LP\* version of the argument, since the force of his argument depends on the fact that John's choice in his example is a SFA, like the businesswoman's, instead of being like the actions of the husband and assassin. But if this is so, then John's situation will also be like the businesswoman's on the account just given of SFAs. Since both of them are simultaneously trying to do *both* of the things they may do (choose to help or go on, overcome the temptation to arrive late or not), they will do either with intent or on purpose, as a result of wanting and trying to do it—that is, intentionally and voluntarily. Thus, their “failing” to do one of the options will not be a mistake or accident, but a voluntary and intentional doing *of the other*.

Likewise, businesswoman\* and John\* are simultaneously trying to do both things in their respective worlds; and they will not “fail” to act on moral or weak-willed motives by mistake or accident, as the case may be, but by voluntarily and intentionally choosing to act on the opposing motives. The point is that in self-formation of these kinds (SFAs), failing is never *just* failing; it is always also a *succeeding* in doing something else we wanted and were trying to do. And we found that one can be responsible for succeeding in doing what one was trying to do, even in the presence of indeterminism. So even if we add the LP\* requirement of more-than-one-way voluntariness and intentionality to the argument of (a)-(e), the argument remains invalid for cases like the businesswoman's and other SFAs, like John's.

But one might argue further, as Mele does, that John and John\* (and businesswoman and businesswoman\*) not only had the same capacities, motives, characters, and the like prior to choice, but they made exactly the same *efforts* as well. And this does seem to suggest that the success of one and failure of the other was a matter of mere luck or chance, so that John and the businesswoman were not responsible. But again the inference is too hasty. Note, first, that husband and husband\* also made the same efforts (as well as having the same capacities, motives, and characters) up to the very moment of breaking of the table. Yet it does not follow that the husband is not responsible when he succeeds. And *both* the businesswoman and businesswoman\*, and John and John\*, are in the position of the husband in their respective worlds, since both will have succeeded in doing what they were trying to do.

But one may still want to object: if the businesswoman and businesswoman\*, and John and John\*, make exactly the same efforts, how can it *not* be a matter of chance that one succeeds and the other does not, in a way that makes them not responsible? To which I reply: But if they both succeeded in doing what they were trying to do (because they were simultaneously trying to do both things), and then having succeeded, they both *endorsed* the outcomes of their respective efforts (that is, their choices) as what they were trying to do, instead of disowning or disassociating from those choices, how then can we *not* hold them responsible? It just does not follow that, because they made exactly the same efforts, they chose *by chance*.

To say something was done “by chance” usually means (as in the assassin and husband cases when they fail), it was done “by mistake” or “accidentally,” “inadvertently,” “involuntarily,” or “as an unintended fluke.” But none of these things holds of the businesswoman and John either way they choose. Unlike husband\*, businesswoman\* and John do not fail to overcome temptation by mistake or accident, inadvertently or involuntarily. They consciously and willingly fail to overcome temptation *by* consciously and willingly choosing to act in selfish or weak-willed ways. So, just as it would have been a poor excuse for the husband to say to his wife when the table broke that “Luck or chance did it, not me,” it would be a poor excuse for businesswoman\* and John to say “Luck or chance did it, not me” when they failed to help the assault victim or failed to arrive on time.

Worth highlighting in this argument is the point that we cannot simply say the businesswoman and businesswoman\* (or John and John\*) made exactly the same *effort* (in the singular) in their respective possible worlds and one succeeded while the other failed. We must say they made exactly the same *efforts* (plural) in their respective worlds. Mentioning only one effort prejudices the case, for it suggests that the failure of that effort in one of the worlds was a *mere* mistake or accident, when the fact is that both of the agents (P and P\*) made *both* efforts in *both* worlds. In one world, one of the efforts issued in a choice and in the other world, a different effort issued in a different choice; but neither was merely accidental or inadvertent in either world. I would go even further and say that we may also doubt that the efforts they were both making really were exactly the same. Where events are indeterminate, as are the efforts they were making, there is no such thing as exact sameness or difference of events in different possible worlds. Their efforts were not exactly the

same, nor were they exactly different, because they were not exact. They were simply unique.<sup>23</sup>

One might try another line: perhaps we are begging the question in assuming that the outcomes of the efforts of the businesswoman and her counterpart were *choices* at all. If they were not choices to begin with, they could not have been voluntary choices. One might argue this on the grounds that (A) "If an event is undetermined, it must be something that merely happens and cannot be somebody's choice"; and (B) "If an event is undetermined, it must be something that merely happens, it cannot be something an agent does (it cannot be an action)." But to see how question-begging these assumptions are, one has only to note that (A) and (B) imply respectively (A') "If an event is a choice, it must be determined" ("All choices are determined") and (B') "If an event is an action, it must be determined" ("All actions are determined"). Are these supposed to be a priori or analytic truths? If so, then long-standing issues about freedom and determinism would be settled by fiat. If an event were not determined, it could not be a choice or action necessarily or by definition.<sup>24</sup>

This explains the businesswoman's suspicions when she exited the neuroscientists' offices. They told her that when she "chose" to go back to help the assault victim the day before, there was some indeterminism in her neural processes prior to choice. She accepted this as a correct empirical finding. But she was suspicious when the neuroscientists tried to get her to make the further inference from those findings which she did not really *choose* to help the assault victim yesterday. She refused to accept that conclusion, and rightly so. For in drawing it, they were going beyond their empirical findings and trying to foist on her the a priori assumption that if an event was undetermined, it could not have been her choice or could not have been something she did. She rightly saw that there was nothing in the empirical evidence that required her to say that. To choose is consciously and deliberately to form an intention to do something; and she did that, despite the indeterminism in her neural processes (as did businesswoman\* when she chose to go on to her meeting).

#### VI. FINAL CONSIDERATIONS: CONTROL AND EXPLANATION

But it is one thing to say that she chose and another to say she chose *freely* and *responsibly*. This would require that she not only chose, but had voluntary *control* over her choice either way. We have not talked at length to this point about the matter of control (considerations (1)

<sup>23</sup> See *The Significance of Free Will*, pp. 171-74.

<sup>24</sup> *Ibid.*, pp. 183-86, for a fuller account of why indeterminism does not rule out action or choice.

and (3) of section II) and must now do so. For this may be the reason why we may think the choices made by the businesswoman and businesswoman\* (or John and John\*) could not be responsible, if they were undetermined. We might deny that they had voluntary control over what they chose, where voluntary control means being able to bring about something in accordance with one's will or purposes (or, as we often say, the ability to bring something about "at will").

One thing does seem to be true about control which critics of indeterminist freedom have always maintained: indeterminism, wherever it appears, does seem to *diminish* rather than enhance agents' voluntary control (consideration (3) of section II). The assassin's voluntary control over whether or not the prime minister is killed (his ability to realize his purpose or what he is trying to do) is diminished by the undetermined impulses in his arm—and so also for the husband and his breaking the table. Moreover, this limitation is connected to another, which I think we must also grant—that indeterminism, wherever it occurs, functions as a *hindrance* or *obstacle* to our purposes that must be overcome by effort (consideration (3)).

But recall that in the businesswoman's case (and for SFAs generally, like John's), the indeterminism that is admittedly diminishing her ability to overcome selfish temptation, and *is* indeed a hindrance to her doing so, is coming from her own will—from her desire and effort to do the opposite—since she is simultaneously trying to realize two conflicting purposes at once. Similarly, her ability to overcome moral qualms is diminished by the fact that she also simultaneously wants and is trying to act on moral reasons. If we could look at each of the two competing neural networks involved separately, abstracting from the other, the situation would look analogous to the situations of the husband and the assassin. The agent would be trying to do something while being hindered by indeterminism coming from an external source. But, in fact, we cannot look at the two networks separately in this way because, in reality, they are connected and interacting. The indeterminism that is a hindrance to her fulfilling one is coming from its interactions with the other. The indeterminism, therefore, does not have an external source. It is internal to her will, and hence to her self, since she identifies with both networks and will identify with the choice reached by either of them as her choice.

The upshot is that, despite the businesswoman's diminished control over *each* option considered separately, due to a conflict in her will, she nonetheless has what I call *plural voluntary control* over the two options considered *as a set* (*ibid.*, pp. 134–43). Having plural voluntary control over a set of options means being able to bring about

*whichever* of the options you will or most want, *when* you will to do so, for the reasons you will to do so, without being coerced or compelled in doing so. And the businesswoman (or John) has this power, because whichever of the options she chooses (to help the victim or go on to her meeting) will be *endorsed* by her as what she wills or most wants to do at the moment when she chooses it (though not necessarily beforehand); she will choose it for the reasons she most wants to act on then and there (moral or selfish reasons, as the case may be); she need not have been coerced by anyone else into choosing one rather than the other; and she will not be choosing either compulsively, since neither choice is such that she could not have chosen it then and there, even if she most wanted to.<sup>25</sup>

One must add, of course, that such plural voluntary control is not the same as what may be called *antecedent determining control*—the ability to determine or guarantee which of a set of options will occur *before* it occurs (*ibid.*, p. 144). With respect to undetermined self-forming choices (SFAs), agents cannot determine or guarantee which choice outcome will occur *beforehand*; for that could only be done by predetermining the outcome. But it does not follow that, because one cannot determine which of a set of outcomes will occur before it occurs, one does not determine which of them occurs *when* it occurs. When the conditions of plural voluntary control are satisfied, agents exercise control over their present and future lives then and there by deciding.

But can we not at least say that, if indeterminism is involved, then *which* option is chosen is “arbitrary”? I grant that there is a sense in which this is true. An ultimate arbitrariness remains in all undetermined SFAs because there cannot, in principle, be sufficient or overriding *prior* reasons for making one set of competing reasons prevail over the other. But I argue that such arbitrariness relative to prior reasons tells us something important about free will. It tells us, as I have elsewhere expressed it, that every undetermined self-forming choice (SFA) “is the initiation of a ‘value experiment’ whose justification lies in the *future* and is not fully explained by the *past*. [Making such a choice], we say in effect, ‘Let’s try this. It is not required by my past, but is consistent with my past and is one branching pathway my life could now meaningfully take. I am willing to take responsibility for it one way or the other’ ” (*ibid.*, pp. 145-46). To initiate and take responsibility for such value experiments whose justification lies in the future, is to “take chances” without prior guarantees of success. Genuine self-formation requires this sort of risk-taking and indeterminism is a part of it. If there are persons who need to

<sup>25</sup> *Ibid.*, pp. 133-38, where a more detailed case is made for each of these claims.

be certain in advance just exactly what is the best or right thing to do in every circumstance (perhaps to be told so by some human or divine authority), then free will is not for them.

This point also throws light on why the luck argument fails, even in the stronger LP\* version, despite its initial plausibility. Consider the move from step (c)—the agents *P* and *P*\* have the same powers, characters, motives, and the like, prior to *t* in the two possible worlds—to step (d), which says it was a matter of luck or chance that *P* did *A* and *P*\* did *B* at *t*. An important reason given for this move was that, if both agents have all the same prior powers, characters, motives, and the like, there can be no “explanation of the difference in choice” between the two agents in terms of their prior reasons or motives; and this is taken to imply that the difference in choices in the two worlds is a matter of luck or chance *in a way* that precludes responsibility.

But this move, like others discussed earlier, is too hasty. The absence of an explanation of the difference in choice in terms of prior reasons does not have the tight connection to issues of responsibility one might initially credit it with. For one thing, the absence of such an explanation does not imply (as I have been arguing throughout this paper) that businesswoman and businesswoman\* (John and John\*) (1) did not *choose* at all, nor does it imply that they did not both choose (2) *as a result of their efforts*, nor that they did not choose (3) *for reasons* (different reasons, of course) that (4) they most wanted to choose for *when* they chose, nor that they did not choose for those reasons (5) *knowingly* and (6) *on purpose* when they chose, and hence (7) *rationally*, (8) *voluntarily*, and (9) *intentionally*. None of these conditions is precluded by the absence of an explanation of the difference of choice in terms of prior reasons. Yet these are precisely the kinds of conditions we look for when deciding whether or not persons are responsible.

I suggest that the reason why these conditions are not excluded is that the explanation of the difference of choice in the two possible worlds which is missing is an explanation in terms of *sufficient* or *conclusive* reasons—one that would render an alternative choice, given the same prior reasons, irrational or inexplicable. And, of course, *that* sort of explanation is not possible for undetermined SFAs, when there is conflict in the will and the agent has good (but not decisive or conclusive) prior reasons for going either way. But neither is that sort of explanation required to say that an agent acts as the result of her effort for reasons she most wants to act on then and there. In sum, *you can choose responsibly for prior reasons that were not conclusive or decisive prior to your choosing for them.*

I said a moment ago that such arbitrariness relative to prior reasons tells us something important about free will—that every self-forming choice is the initiation of a value experiment whose justification lies in the future and cannot be fully explained by the past. It is worth adding in this regard that the term ‘arbitrary’ comes from the Latin *arbitrium*, which means ‘judgment’—as in *liberum arbitrium voluntatis* (“free judgment of the will”)—the medieval designation for free will. 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 react. The author must make a “judgment” (*arbitrium*) about how she will react that is not determined by the heroine’s already formed past, which does not give unique direction. In this sense, the author’s judgment of how she will act is “arbitrary,” but not entirely so. It has input from the heroine’s fictional past and, in turn, gives input to her projected future.

In a similar manner, agents who exercise free will are both authors of, and characters in, their own stories at once. By virtue of “self-forming” judgments of the will (*arbitria voluntatis*), they are “arbiters” of their own lives, taking responsibility for “making themselves” out of past that, if they are truly free, does not limit their future pathways to one. If someone should charge them with not having a sufficient or conclusive prior reason for choosing as they did, they may reply as follows: “Perhaps so. But that does not mean I did not *choose*, and it does not mean I did not choose for *good* reasons, which I stand by and for which I take responsibility. If I lacked sufficient or conclusive prior reasons, that is because, like the heroine of the novel, I was not a fully formed person before I chose—and still am not, for that matter.”<sup>26</sup> Like the author of the novel, I am in the process of writing a story and forming a person (who, in my case, is myself). It is a heavy burden, but an eminently human one.”

ROBERT KANE

University of Texas/Austin

<sup>26</sup> Jan Branson (in “Alternatives of Oneself,” *Philosophy and Phenomenological Research* (forthcoming)) has made an important distinction that is relevant here—between choosing “alternatives *for* oneself” and choosing “alternatives *of* oneself.” Branson notes that some choices in life are for different courses of action that will make a difference in what sort of person the chooser will become in future. In such cases, agents are not merely choosing alternatives for themselves but are choosing alternatives of themselves. Many SFAs, as I understand them, would be of this kind.