

Suggested Reading

Dualist views of free will are defended by John Eccles and Karl Popper, *The Self and Its Brain* (Springer-Verlag, 1977); Richard Swinburne, *The Evolution of the Soul* (Oxford: Clarendon, 1986); John Foster, *The Immaterial Self* (Routledge, 1991); and J. P. Moreland and Scott Rae, *Body and Soul* (InterVarsity, 2000). Kant's view of free will presented in this chapter appears in his *Critique of Pure Reason* and *Foundations of the Metaphysics of Morals*. Chisholm's agent-causal view as expressed in "Human Freedom and the Self" (appears in several edited volumes: Gary Watson, *Free Will*, 2nd ed. (Oxford, 2003); Robert Kane, *Free Will* (Blackwell, 2002); and Laura Waddell Ekstrom, *Agency and Responsibility: Essays on the Metaphysics of Freedom* (Westview, 2000). Thomas Reid's agent-causal view is sympathetically examined in William Rowe's *Thomas Reid on Freedom and Morality* (Cornell, 1991).



Actions, Reasons, and Causes

1. Simple Indeterminism

Some modern libertarians argue that libertarian free will can be explained without the need to appeal to "extra factors" of the kinds discussed in the preceding chapter, such as minds outside space and time or non-event agent-causation. One theory that takes this line is called *simple indeterminism*. The key to understanding free will, according to simple indeterminists, is a distinction between two ways of *explaining* events—explanations in terms of *causes* and explanations in terms of *reasons* or *purposes*. Free actions are *uncaused* events, according to simple indeterminists, but the fact that free actions are uncaused does not mean they occur merely by chance or randomly. The occurrence of free actions, though uncaused, can be explained in terms of the reasons and purposes of agents.

Understanding this simple indeterminist view requires discussion of two topics that play an important role in debates about free will but have not to this point received enough attention: (1) the nature of *explanation* and (2) the nature of *action*. Many problems about free will discussed in chapters 4 and 5 concern the question of how free actions can be *explained* if they are undetermined or uncaused. Questions about how free actions can be explained in turn lead to deeper questions about what makes something an *action* in the first place rather than an event that merely happens (say, by chance or accident). We must now consider these questions about the nature of explanation and action.

An *explanation* of any kind is an answer to a *why* question: Why does something exist? Why did it occur? Why is it so? But in the case of events,

there are two kinds of answers to the question "Why did it occur?"—an explanation in terms of causes (e.g., the fire was caused by an explosion) and an explanation in terms of reasons and purposes. And the explanations we usually give when talking about human *actions* are explanations in terms of reasons and purposes. For example, when we ask "Why did Mary enter the room?" we give her reasons in the form of her wants, desires, beliefs, intentions, and goals. Mary entered the room because she *wanted* to find her keys, *believed* she may have left the keys there, and had the *purpose* or *goal* of finding them. Citing these reasons and purposes explains *why* Mary acted as she did. But it does not follow that Mary was caused or determined to act that way, say simple indeterminists. For reasons and purposes are not causes of action, according to them; and explanations in terms of reasons are not causal explanations. Free actions may therefore be *uncaused* without occurring merely by chance or randomly. They occur for a reason or purpose.

But if free actions really are uncaused events, as simple indeterminists claim, what makes them "acts" or "actions" in the first place rather than mere "happenings" occurring out of the blue? (This was the second question just mentioned, about the nature of *action*.) One prominent simple indeterminist, Carl Ginet, answers this question by arguing that an action, such as Mary's entering the room, begins with a simple mental act, a *volition* or act of will that initiates the action. What makes this volition and the action initiated by it actions rather than things that merely "happen" to Mary, according to Ginet, is that the volition and action have a certain "actish phenomenal quality"—that is, the volition and the action are directly *experienced* by Mary as something she is doing rather than something that happens to her.¹

We are all aware of this difference in things that occur in our minds. Some mental events, such as the sudden occurrence of a thought or memory or image, seem to merely come *upon* us or happen *to* us in a way that is not under our control. But other mental events, like concentrating in the attempt to solve a problem, or making a decision, are things we seem to be doing that are under our control. Mental events of the latter kind, those that seem to be under our control, according to Ginet, have this "actish phenomenal quality"; and it is the presence of this experienced quality that makes them actions rather than mere happenings. Of course, not all actions are *free* actions. Ginet's actish phenomenal quality guarantees only that something is an action. For an action to be *free*, he insists, it must not only have this actish phenomenal quality, it must be done for a reason or purpose *and* it must be undetermined.

2. *Objections to Simple Indeterminism*

Many philosophers question the simple indeterminists' claim that reasons for actions are not causes of actions. Mary's reasons for entering the room were that she *desired* to find her keys and *believed* she might have left them in the room she entered. Citing these reasons explains why Mary entered the room, to be sure. But why can't we also say that her having this desire and belief were among the causes of her entering the room—and that is *why* they explain her behavior? Maybe the desire and belief were not the *sole* causes of Mary's action; and perhaps they did not determine that she would enter the room. Our reasons may "incline without necessitating," as Leibniz said. They may make it more likely that we will act in certain ways. But when we do act in these ways, it is natural to say that our desires and beliefs causally influence our acting, even if they do not determine it.

By comparison, a crack in a bridge support may make a bridge collapse more likely. The crack alone will not cause the collapse in the absence of a strong wind. Yet if the bridge does collapse in a strong wind, the crack in the support will have been one of the causes of the failure. So it would also be with desires, beliefs, and other reasons for action, say these critics of simple indeterminism. When we do act on them, they are among the causes of our actions, though not necessarily the sole causes.

In response to this objection, simple indeterminists, such as Ginet, concede that desires, beliefs, and other reasons do *influence* actions, but not by causing them. To understand how desires and other reasons might influence actions without causing them, one must bring in two other notions that are important in free will debates—the notions of *intention* and *purpose*. Free actions are actions we do *intentionally* or on *purpose*, not by accident or mistake. Mary's action was intentional, not accidental. When she entered the room, she *intended* to find her keys. Her purpose was therefore "to find her keys." An *intention* is a state of mind; and what we call a *purpose* is the mental *content* of the intention—what the intention is *about*. Thus, if I am walking to the store and have in my mind "the *intention* <to buy a jacket>" then my *purpose* is "<to buy a jacket>"—what my intention is an intention *to do*.

Ginet now adds that desires and other reasons influence actions, not by causing them, but by entering into the *contents* of our intentions to perform the actions. Thus, Mary's desire to find her keys influenced her entering the room because she *intended* <to enter the room *in order to satisfy the desire* to find her keys>. Reference to the desire is included in the purpose (which is signified by the brackets). In this way, Mary's *intention* and

purpose provide the explanatory link between the *action* (entering the room) and her *desire* (to find the keys). The desire influences the action, not by causing it, but by being referred to in the intention to perform the action. But does this *intention* itself cause the action? No, says Ginet. The intention *explains* the action not by causing it, but simply by *referring* to the action (it is the intention <to enter the room>) and by linking the action to the reason (<to satisfy the desire>). Thus Mary's acting can be explained and is not merely arbitrary, *even though it was undetermined*.

Critics object, however, that many of our reasons for acting never explicitly enter our intentions in the way Ginet describes, yet they still influence our actions. Freud and other psychoanalysts have made us aware, for example, that many of our desires and other reasons for acting are *unconscious* reasons. In addition, we often *repress* the real reasons for our actions or deceive ourselves about why we are doing something. Suppose Mary's real reason for entering the room was to wake up her brother, who was sleeping there, though she repressed that reason and deceived herself into thinking she was entering the room to find the keys. (In fact, the keys were more likely to have been in another room.) Since childhood, Mary had always resented the fact that her brother was an earlier riser and out of meanness woke her up on school days before she wanted to be wakened. In such a case, it is natural to say that wanting to wake her brother was a cause of Mary's entering the room even though it was not the reason referred to in her intention. There are many reasons (wants, desires, beliefs, preferences, aversions, likes, dislikes, etc.)—both conscious and unconscious—that influence our acting as we do. It is not credible, as Alfred Mele points out in his book *Motivation and Agency*, that all these reasons must be referred to in the contents of our intentions in order to influence our actions.² It is more natural to think that reasons can causally influence our actions even if they do not explicitly enter into our intentions.

A second related objection to simple indeterminism concerns Ginet's claim that volitions and other actions are distinguished from things that merely happen *to us* by an "actish phenomenal quality." This means we directly experience our actions as things we are doing rather than things that are happening to us. But could this experience be *illusory*? If our free actions really are uncaused, might we be experiencing them *as if* they were our actions when they really are not. One critic of simple indeterminism, R. E. Hobart, puts this objection in the following way:

In proportion as an act of volition starts of itself without cause, it is exactly, so far as the freedom of the individual is concerned, as if it had been thrown into the mind from without—"suggested to him by a freakish demon."³

Another critic of simple indeterminism, Timothy O'Connor, puts the objection this way:

The fact that free actions have uncaused volitions at their core is *prima facie* puzzling. If [a volition] is uncaused, if it is in no sense determined to occur by anything at all, then it is not determined to occur by me in particular. And if I don't determine it, then it's not under my control.⁴

3. Agent-causation Revisited

O'Connor argues that simple indeterminism is inadequate at this point unless we add to it a notion of non-event *agent-causation* like that of Chisholm and Reid discussed in chapter 5. Free actions may be uncaused *by prior events*, O'Connor says, but they cannot be uncaused by *anything*. If a free action was "uncaused . . . by anything," then it would not be caused "to occur *by me* in particular" and would not be "under my control." O'Connor does agree with simple indeterminists that explanations of actions in terms of *reasons* are not explanations in terms of *causes*. He also accepts Ginet's idea that desires and other reasons can explain actions by referring to the agent's intentions. Thus O'Connor agrees that we can explain why Mary entered the room by saying she had the intention <to satisfy her desire to find her keys>.

But O'Connor thinks we must also ask where this *intention* of Mary's came from. If Mary's intention to enter the room to satisfy the desire was not caused by her desire or other reasons, what caused it? This is where O'Connor thinks a notion of non-event agent-causation like that of Chisholm and Reid must be brought in. Mary's intention to enter the room was not caused by her desire or any other reasons and was not determined by any prior events. But the intention was nonetheless directly caused by the agent, Mary, herself; and it was caused by her in a special way that cannot be explained in terms of causation by prior events. In short, we must invoke what Chisholm called *immanent* or direct causation of events or states by agents rather than the *transeunt* causation of events by other events.

Simple indeterminists, such as Ginet, are suspicious of this addition of a special kind of non-event agent-causation. They think it is unnecessary to "complicate our picture of free agency" with this additional notion. Another simple indeterminist, Stewart Goetz, states this objection to agent-causation in the following way. Goetz says that, on his simple indeterminist view, a choice—such as Mary's choosing to enter the room—is an uncaused event that is directly under the control of the agent.⁵ If Mary

did not have direct control over her choice, says Goetz, it would not *be* her choice. O'Connor's response is that Goetz is getting this result "for free" by simply *defining* a choice as an event that is (a) uncaused and therefore undetermined and yet (b) under the control of the agent. The problem, according to O'Connor, is to explain *how* an event could be uncaused by prior events and yet under the control of the agent.

"But what is O'Connor's alternative?" asks Goetz. It amounts to interpreting a free choice as the agent-causing-of-an-intention in a special non-occurrent way and then *defining* this special relation of agent-causation so that it is (a) essentially undetermined and (b) also essentially under the control of the agent. Goetz then adds: If I am getting my result "for free," then agent-causalists, such as O'Connor, are getting their result for free as well; and they are adding an extra and obscure notion of non-event causation to do it. If it is illegitimate, Goetz asks, for the simple indeterminist to define *Mary's choosing* (to enter the room) as essentially an exercise of power that is uncaused by prior events, yet under the direct control of the agent, then why isn't it just as illegitimate for the agent-causalist to define *Mary's agent-causing her intention* (to enter the room) as an exercise of power that is uncaused by prior events, yet under the direct control of the agent?

This is a potent question. Compatibilists, such as Watson, are likely to say at this point that *both* parties—simple indeterminists and agent-causalists—are getting their results illegitimately: by definition or stipulation. But O'Connor has a response to Goetz's objection. He insists that the agent-causalist *is* adding something important. By interpreting *Mary's choosing* as *Mary's agent-causing-her-intention*, the agent-causalist is bringing out the fact that choices are not "simple" mental events, as Goetz and other simple indeterminists claim. Choices have a causal structure. A choice to do something is an *agent's-bringing-about-or-causing-an-intention* to do it. By thus noting that free choices are agent-causings and not simple events, O'Connor argues, agent-causalists, unlike simple indeterminists, can explain *why* free choices are essentially uncaused by prior events.⁶

To explain this, O'Connor asks us to consider that ordinary causation by events has the following structure: Event *e'* (e.g., the lighting of a match) causes event *e''* (an explosion). He then argues that causal relations between events like this (*e'* causes *e''*) cannot themselves be caused—at least not directly. We *can* say that the *striking* of the match (*e*) caused the match's *lighting* (*e'*) to cause the *explosion* (*e''*). But in that case we are saying that event *e* (the striking) causes the *first event* in the causal relation, namely *e'* (the lighting of the match), and *then e'* causes the second event, *e''* (the explosion). In other words, O'Connor argues, a causal

relation between events (such as *e'* causes *e''*) can only be caused indirectly by causing the first event (*e'*) in the causal relation, which then causes the second event (*e''*).⁷

But in the case of *agent-causation*, he argues, the causal relation does not have the usual form of causation between events (*e* causes *e'*). Instead, agent-causation has the form "A causes *e*," where the first term is not an event at all, but an agent, an enduring substance. And, O'Connor argues, "there appears to be no way of getting a grip on the notion of an event of *this sort*" (agent A causes event *e*) having a sufficient cause. "Because of its peculiar causal structure [*A causes e*], there is no event at its front end, so to speak [that could be caused by some other event] but only an enduring agent."⁸ So an agent-causal relation cannot in principle itself be caused by other events. By adding such a notion, agent-causalists can explain, as simple indeterminists cannot, *why* free choices cannot be determined.

4. Actions and Events

One difficulty with the preceding argument concerns the nature of *action*. O'Connor is bringing out something important when he says that choices are not simple events. They appear to have a causal structure. A *choice* to do something (such as enter a room) is an agent's-bringing-about-or-causing-the-intention to do that thing. But the problem is that something similar could be said about *actions* of many kinds, not merely choices. To *act*, in general, is to *bring about or cause some event or state of affairs*. For example, to *kill the king* is to bring about (or cause it to be the case) that *the king is dead*. To *raise your arm* is to bring about (or cause it to be the case) that *your arm goes up*. To *turn on the light* is to bring it about that *the light is on*, and likewise for other actions.

This is the feature that makes *actions* different from simple *events* or happenings. Actions have the form "Agent (A) brings about or causes an event or state (*e*)," where the first term of the causal relation is an agent and the second term is an event or state of affairs. This feature of actions is one of the things that lends plausibility to agent-causal theories. But this feature of actions also raises questions about O'Connor's argument. For, if it is true that a causal relation of the form "A causes *e*" cannot itself be causally determined by prior events *because its first term is an agent and not an event*, then this would be true of actions, in general, not merely of free actions. For actions, in general, have this agent-causal form. That is what distinguishes them from mere events. If the argument worked, it would show that for something to be an action, whether free or unfree, it could not in principle be determined.

Some people might want to accept this strong conclusion. They might say that all actions must of necessity be undetermined. Then, if we lived in a determined world, no one would really *do* anything. Things would merely happen. There would be a “flow of events,” but no real agency. But most people do not want to go that far. Even libertarians and incompatibilists usually insist that it is only *free* actions that must be undetermined, not all actions whatsoever. When persons act compulsively or are forced to do certain things (say, hand over their money when a gun is held to their head), they *do* something, though not freely. In fact, O'Connor himself does not want to say that all actions are essentially undetermined: only free actions are. But then, it is not sufficient for him to argue that a causal relation of the form “A causes e” could not be causally determined because *its first term is an agent* rather than an event. (For, all actions having this agent-causal form need not be undetermined.) He must add that *free* actions are unique because they are agent-caused in the *special* non-event or non-occurrent way that by its nature cannot in principle be determined. This claim, however, goes well beyond, and is not supported by, the claim that free actions have an agent-causal structure (A causes e) alone. One might argue therefore as Goetz does, that this further claim amounts merely to stipulating that free actions involve an agent-causal relation of a special kind that is (a) essentially undetermined and (b) also essentially under the control of the agent.

It is an important fact about actions and choices, to be sure, that they have an agent-causal structure: John's raising his arm is bringing it about (or causing it to be the case) that his arm goes up; Mary's making a choice is bringing it about (or causing it to be the case) that she has an intention or purpose to do something. Agent-causalists, such as Chisholm and O'Connor, correctly draw our attention to this fact. But having such an agent-causal structure does not *alone* prove that actions or choices cannot in principle be caused or determined by prior events. Stronger arguments are needed to show that.

5. *The Causal Theory of Action*

This debate about the causal structure of action is related to another feature of the simple indeterminists' view discussed in section 1, namely, the claim that *reasons* for actions are not *causes* of actions. As noted, many philosophers question the simple indeterminists' claim that reasons cannot be causes. Mary's reasons for entering the room were that she *desired* to find her keys and *believed* she might have left them in the room. Citing these reasons explains why Mary entered the room. But why, these

philosophers ask, can't we also say that Mary's having this desire and belief were among the causes of her entering the room? The desire and belief need not have been the sole causes of Mary's action, just as the structural defect in the bridge was not the sole cause of the bridge's collapse. But we could still say that Mary's desires, beliefs, and other motives were among the causes of her action.

Philosophers who take this line—who insist that desires, beliefs, and other reasons are causes of action—are often called *causal theorists of action*.⁹ Causal theorists of action agree with agent-causalists that actions have an agent-causal structure: they agree that an action is an agent's-bringing-about-or-causing-something to occur. But (in opposition to agent-causalists) causal theorists of action argue that the agent-causal structure of action *can* be explained in terms of causation by prior *events* or *states of affairs*. Mary's entering the room, they say, was caused by her *intention* to enter the room; and her intention was caused by her *choice* to enter the room; and her choice to enter the room was caused by her *desire* to find her keys and by her *belief* that her keys might be in the room. To explain actions, according to causal theorists, one does not have to postulate any additional form of non-event agent-causation over and above causation by mental states and processes, such as beliefs, wants, desires, and intentions. This is true of *choices* as well as of actions of other kinds, according to causal theorists: Mary's *choice* to enter the room was also caused by her desires and beliefs, together with other mental events, such as her memories and perceptions, that entered into her deliberation and, through her deliberation, causally influenced the choice she made.

As you might guess, many causal theorists of action tend to be compatibilists or even determinists about free will. They reason that, if choices and actions can be caused by the agent's reasons and other mental states, then choices and actions might also be *determined* by the agent's reasons and other mental states. In fact, the causal theory of action is often invoked to refute libertarian theories of free will, such as simple indeterminism, which claim that free actions or choices are not caused by reasons and therefore cannot in principle be determined.

6. *Causation and Determinism*

But one can agree with causal theorists that reasons may be causes of action without necessarily being a compatibilist or a determinist about free will. For the fact is that all causes need not be *determining* causes. Some causes are merely probabilistic; they make it more likely that certain events will occur without determining that those events will occur. And

this might be the case with reasons and motives as well. Free choices and actions may be causally influenced by the agent's reasons or motives without being determined by those reasons or motives. As Leibniz said, reasons may "incline without necessitating." Mike's desire to surf along with other reasons may incline him to choose to vacation in Hawaii without necessitating or determining that choice; and his desire to ski along with other reasons may incline him to choose to vacation in Colorado without necessitating that choice.

But we may wonder what "tips the balance," if Mike might choose either Hawaii or Colorado and neither choice is determined by his reasons. Perhaps *this* is the point at which one must introduce some kind of agent-causation "over and above" causation by prior reasons and motives. That is the line taken by another agent-causalist, Randolph Clarke.¹⁰ Clarke is unpersuaded by arguments of simple indeterminists and other agent-causalists, such as O'Connor, that reasons cannot be causes of actions. Clarke thinks that many reasons or motives, conscious and unconscious, may causally influence our actions even though they are not referred to in our intentions. But he still believes that non-event agent-causation is needed to explain what tips the balance between the reasons for one choice or the other when neither set of reasons is determining. Somehow Mike himself (the *agent*) must cause the choice of Hawaii (or Colorado) in a way that cannot be completely explained in terms of his prior reasons or his prior deliberation or in terms of any prior events whatsoever.

But how does appealing to agent-causation explain why Mike's tipping the balance in one way *rather than* the other is not arbitrary or random, since his reasons and motives may have inclined him in either direction? Clarke concedes that introducing non-event agent-causation at this point does not answer puzzles about arbitrariness of this sort concerning libertarian free agency. But introducing agent-causation, he argues, does at least account for the fact that the agent, Mike, has *control* over, and *produces*, the choice that is finally made, as opposed to one set of reasons simply "winning out" over the other set by mere chance. Yet critics of agent-causation, such as Watson, respond that postulating agent-causation at this point does not seem to explain *how* the agent controls or produces one choice rather than the other either.¹¹ The agent-causalist says that the agent controls or produces one outcome rather than the other without really explaining *how* the agent can do this except randomly or arbitrarily. This criticism reminds one of Watson's objection noted in the preceding chapter: that agent-causation merely "labels what libertarians need," rather than explaining it. Clarke might respond that agent-causation *does* nonetheless correctly represent what libertarians need—namely, something to tip the balance.

Ginet and O'Connor have a different objection to Clarke's agent-causal view. They argue that if libertarians concede, as Clarke does, that desires, beliefs, and other reasons can be causes of action (even indeterministic causes), then libertarians risk making agent-causation of a special non-event kind superfluous. Ginet asks: does the agent cause in Clarke's theory, supply some extra "oompf" or force that the reasons and other mental and physical events do not supply, an extra force that tips the balance?¹² Clarke admits that this cannot be what agent-causation adds. We cannot think of non-event agent-causation as some kind of extra *force*, either physical or mental, that tips the balance. To construe agent-causation in such mechanical "push/pull" terms would be to reduce it to another kind of event causation, which it is not.

But, says Ginet, that seems to be the picture we have, if we allow that reasons may be indeterministic or probabilistic causes of free actions. For then reasons would supply *some* of the force inclining us to make a particular choice, but not enough. The extra force or "oomph" would have to be supplied by the agent. Yet that picture cannot be right, Ginet argues, if agent-causation is not just another form of causation by forces and events. A similar criticism is made by O'Connor. He says that an agent-causation that is irreducible to event causation cannot "be fitted into or on top of an unbroken chain of event causation," including causation by reasons, as Clarke suggests. "Once we recognize free will to involve a type of undetermined, direct control" of the kind that non-event agent-causation requires, "we have to reject the completeness of the simple, continuous-flow-of-events picture of nature."¹³

But Clarke responds that such a view of agent-causation would require that agent-causation (and hence free will) must "interrupt" or "disrupt" the ordinary pattern of events in nature and perhaps that it would in some way violate the laws of nature. And this would make agent-causation (and libertarian free will) mysterious or something of a miracle. One possible reply suggested by O'Connor and others is that non-event agent-causation is a special capacity of organisms that *emerges* in nature but is no longer reducible to natural flow-of-events picture of nature.¹⁴ This suggestion would require further development, however, to explain how, if at all, such an emergent capacity would not "interrupt" the ordinary pattern of events in nature or why it would not violate the laws of nature. Perhaps, to make ultimate sense of agent-causation, one might have to revert after all, to the dualistic picture of a mind and body, in which the mind is somehow outside the natural order of events but capable of intervening in the physical world to "tip the balance." Both Clarke and O'Connor would like to avoid a mind-body dualism of this kind, and they do not want to claim that free will must violate natural laws. But their debate makes some philosophers

wonder whether making sense of free will in agent-causationist terms might require a dualist view of mind after all.¹⁵

7. *Deliberation and Causal Indeterminism*

The final libertarian theory I want to consider in this chapter takes a very different approach to explaining libertarian free choices. This view rejects both simple indeterminism and agent-causation. Instead it focuses on the process of deliberation. When we deliberate, for example, about where to vacation or which law firm to join, many different thoughts, images, feelings, memories, imagined scenarios, and other considerations pass through our minds. Deliberation can be quite a complex process. When Mike thinks about Hawaii, he pictures himself surfing, walking on sunny beaches, eating in his favorite Hawaiian restaurants; and these various thoughts incline him to choose Hawaii. But he also thinks about skiing, sitting by a fireplace after a long day on the slopes, and visiting with friends he knows in Colorado; and he leans toward Colorado. Back and forth he goes, until after a period of time considerations on one side outweigh the others and he finally chooses one option. (Unless, of course he is one of those indecisive types who finds it hard to make up his mind.)

In the course of such deliberations—which may sometimes take hours or days and may be interrupted by daily activities—new thoughts, memories or images can often come to mind that influence our deliberations. Mike may suddenly remember a lively nightclub he visited in Honolulu when he was last there—great music, great girls—and the idea of going back to this place gives him an added reason to favor Hawaii, a reason that hadn't previously entered his deliberation. Other images that flit through his mind may turn him against Hawaii. Imagining himself out on the beach all day, suddenly he remembers his doctor's warning about not getting too much sun if he wants to avoid skin cancer.

Now one could imagine that some of these various thoughts, memories, and imagined scenarios that come to mind during our deliberations are undetermined and arise by chance and that some of these "chance selected considerations" might make a difference in how we decide. If this were to happen in Mike's case, the course of his deliberation, hence his choice, would be undetermined and unpredictable. A Laplacian demon could not know in advance which way Mike would go, even if the demon knew all the facts about the universe prior to Mike's deliberation, for these facts would not determine the outcome. Yet Mike would still have control over his choice in a certain sense. He could not control all the thoughts and

imagined scenarios that come to mind by chance. But he would be in control of how he reacted to those thoughts and imaginings once they did occur. And his choice of Hawaii in the end would be perfectly rational, not arbitrary, if the weight of all the considerations that did come to mind (some of them by chance) weighed in favor of Hawaii. In this way, choices could thus be controlled and rational even though indeterminism was involved in the deliberations leading up to them.

A view of this kind is called *causal indeterminism* or *event-causal libertarianism*, for it allows that our thoughts, images, memories, beliefs, desires, and other reasons may be causes of our choices or actions without necessarily determining choices and actions; and yet this view does not postulate any extra kind of agent-causation either. Two philosophers who have suggested causal indeterminist views of this kind (without endorsing them), Daniel Dennett and Alfred Mele, argue that a view of this kind would give libertarians at least some of the important things they demand about free will.¹⁶ Such a view, for example, provides for an "open future," such as we think we have when we exercise free will. We would not have to think that our choices and the future direction of our lives had somehow been decided long before we were born. Nor would it be possible for behavioral engineers to completely control our behavior as in Walden Two or for Laplacian demons to know what we were going to do, if chance considerations might enter our deliberations.

Yet, as Dennett and Mele also admit, a causal indeterminist view of this deliberative kind does not give us everything libertarians have wanted from free will. For Mike does not have complete control over what chance images and other thoughts enter his mind or influence his deliberation. They simply come as they please. Mike *does* have some control *after* the chance considerations have occurred. But then there is no more chance involved. What happens from then on, how he reacts, is *determined* by desires and beliefs he already has. So it appears that he does not have control in the *libertarian* sense of what happens after the chance considerations occur as well. Libertarians require more than this for full responsibility and free will. What they would need for free will is for the agent to be able to control which of the chance events occur rather than merely reacting to them in a determined way once they have occurred.

Yet, as Mele points out, while this causal indeterminist view does not give us all the control and responsibility that libertarians have wanted, it does give us many of the things they crave about free will (an open future, a break in the causal order, etc.). And it is clearly a possible view. Perhaps it could be further developed to give us more; or perhaps this is as much as libertarians can hope for.

Suggested Reading

Carl Ginet's simple indeterminist view is developed in *On Action* (Cambridge, 1990). Other noncausalist views are Hugh McCann (*The Works of Agency: On Human Action, Will and Freedom*, Cornell, 1998) and Stewart C. Goetz (see references in note 5). Timothy O'Connor's agent-causal view can be found in *Persons and Causes: The Metaphysics of Free Will* (Oxford, 2000), and Randolph Clarke's agent-causal view appears in his *Libertarian Accounts of Free Will* (Oxford, 2003). The indeterminist view described in the final section is developed in Daniel Dennett, "On Giving Libertarians What They Say They Want" (in *Brainstorms*, MIT, 1978) and by Alfred Mele in *Autonomous Agents: From Self-Control to Autonomy* (Oxford, 1995). A different view of this causal indeterminist kind is defended by Laura Waddell Ekstrom *Free Will: A Philosophical Study* (Westview, 2000). *Agents, Causes, and Events: Essays on Free Will and Indeterminism* (Oxford, 1995) edited by O'Connor is a collection of readings for and against the different libertarian views discussed in this chapter. Two other libertarian theories that do not fit clearly into one or another of the categories of this chapter can be found in James S. Felt, *Making Sense of Your Freedom* (Cornell, 1994), and T. L. Pink, *Free Will: A Short Introduction* (Oxford, 2004). Still other libertarian views are mentioned in the readings suggested at the end of chapter 12.



Is Free Will Possible? Hard Determinists and Other Skeptics

1. Oklahoma City and Columbine

On April 15, 1995, a young man named Timothy McVeigh parked a truck loaded with explosives outside a federal office building in Oklahoma City, Oklahoma. The truck exploded, ripping off the front of the building, killing over 130 people, and injuring many others, including office workers, visiting citizens, and federal employees' young children in a day care center in the basement. Why did he do it?

Tim McVeigh had a fairly normal American upbringing in a midwestern town. He joined the army after high school and liked military life so much that he applied for the elite Special Forces. Then things started to turn bad. He was turned down by the prestigious unit, perhaps because of suspicions about his mental stability. This rejection was a bitter disappointment to a sensitive young man, and McVeigh eventually left the military in a state of frustration and resentment. Outside the military, his resentments were further fueled by association with antigovernment militia types and by reading fictional works that described revolts against the U.S. government initiated by bombings of federal buildings. Thus began a downward spiral that led him to allegedly plan and carry out the bombing of the Alfred P. Murrah Building in Oklahoma City.

These are the surface facts. They leave out the fact that McVeigh had help from others, though a wider conspiracy was never proven. But few doubt that he himself was involved. The surface facts also do not tell us what was going on in Tim McVeigh's mind, what demons were haunting him. They do not tell us about his early childhood experiences, or other

factors that may have led him to contemplate and commit such a horrendous act. When most people think about free will in a case like this—when they wonder whether McVeigh was responsible for the act of which he was found guilty—they tend to have the following thoughts. It is understandable that he was disappointed and resentful because he was turned down for Special Forces. But many other young men have been turned down for this elite service and they did not become mass murderers.

Other people also have resentments against the government. But few join militia groups, and most who do join such groups do not actually commit violent acts, much less murder. No, it was said, McVeigh did what he did of his own free will. Others in the same circumstances and with the same experiences would not necessarily have done what he did. We all have difficulties in life, but we have the free choice to make the best of them or the worst. There is such a thing as moral evil; and people like McVeigh are responsible for choosing evil over good. The jury in McVeigh's trial obviously reasoned in this way. McVeigh was given the death penalty and was executed in 2001.

People reasoned similarly about the terrible massacre at Columbine High School in Colorado on April 20, 2000. Two young men, Eric Harris and Dylan Klebold, entered the school with an arsenal of weapons, killing fourteen fellow students and a teacher and injuring many others before turning the guns on themselves. Like McVeigh, Harris and Klebold harbored resentments—in their case because they were constantly ridiculed by classmates and treated as outsiders by most of their peers. Well, one might say, many teenagers are treated that way in high school without turning into mass murderers.

Harris and Klebold were also deeply influenced by violent films and video games. There was a lot of public debate in the press and on TV at the time about the effects of violence in the media and of violent video games on young people. But it was also said that most young people are subjected to violence in the media today and play these games from early ages, yet do not turn into killers like Harris and Klebold. Harris and Klebold were also obsessed with celebrity and wanted to be famous. Obsession with celebrity is another troubling trend among the young (and old) in modern society, but most people do not kill for it. No, it was said, these young men were evil and chose as they did of their own free wills. If Harris and Klebold had not killed themselves, it is not difficult to imagine a jury reasoning in this way and perhaps sentencing them to death.

But there is another way of thinking about these well-known cases, a way favored by *hard determinists*. Hard determinists believe that if you look more deeply into the psychological and other springs of action, you will see that all of us are determined to do what we do, whether it be good or evil; and so none of us is ultimately responsible. People are making

a fundamental mistake, say hard determinists, when they reason that McVeigh, Harris, and Klebold must have acted of their own free wills because other persons in the same circumstances and with the same experiences would not have done what they did. For, no one ever is in exactly the *same* circumstances as anyone else. We all bring different backgrounds, histories, experiences, and temperaments to every situation; and it is naïve to think that people have free will simply because they act differently in *similar* circumstances. If we knew enough about their pasts to really *explain* why McVeigh, Harris, and Klebold did what they did, we would see that any persons who were exactly like them (not merely similar) would have acted as they did in these circumstances. If this were not true, we would not be able to truly explain *why* they did what they did *rather than* something else.

2. Hard Determinism

Such is the view of hard determinism, the third traditional position on free will. At the beginning of chapter 4, I noted that those who believe that free will and determinism are incompatible may take either of two opposing positions. They may deny determinism and affirm free will, as libertarians do. Or they may affirm determinism and deny free will, which is what hard determinists do. Hard determinism can also be distinguished from “soft” determinism, which was defined at the end of chapter 2. Both hard and soft determinists believe in determinism. But soft determinists are *compatibilists* who insist that determinism does not undermine any free will worth having, while hard determinists are *incompatibilists* who take a “harder” line: Since determinism is true, free will does not exist in the true sense required for genuine responsibility, blameworthiness, and desert for deeds and accomplishments. These traditional positions can be nicely summarized in figure 7.1, which returns us to the picture of Incompatibilist Mountain of chapter 4.

Compatibilists and *soft determinists* say you cannot get *up* Incompatibilist Mountain because you cannot show that free will and determinism are incompatible. Soft determinists add that you cannot get down either—you cannot show that an indeterminist free will exists—because

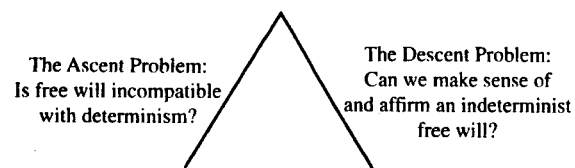


Figure 7.1 Incompatibilist Mountain and the Libertarian Dilemma

determinism is true. (Most other compatibilists also think you cannot get down Incompatibilist Mountain because they do not think an indeterminist free will makes sense.)

Libertarians and *hard determinists*, by contrast, say you *can* get up Incompatibilist Mountain—it can be shown that free will and determinism are incompatible. But hard determinists, in contrast to libertarians, say you cannot get back down because determinism is true. It is cold up there on Incompatibilist Mountain; and hard determinism is a cold view, according to most people, since it requires us to live without free will.

It is not surprising that few thinkers have been willing to embrace such a hard determinist position unqualifiedly, since it seems to require major changes in the way we think about human relations and attitudes, how we treat criminals and assess criminal behavior, and so on. This has not prevented hard determinism from being endorsed by some thinkers, such as Baron d'Holbach in the eighteenth century and Paul Edwards in the twentieth. The controversial American attorney Clarence Darrow was even known for defending hard determinism in the courtroom. Darrow gained fame in the 1931 Scopes trial, in which he defended a Tennessee high school teacher who had been fired for teaching the theory of evolution. But in other cases, such as the equally famous Leopold and Loeb trial, Darrow argued that his clients, Nathan Leopold and Richard Loeb, were not ultimately responsible for doing what they did—for murdering a young boy in cold blood for the sheer pleasure of it—because they were determined to do what they did by their formative circumstances. Few thinkers have been willing to go as far as Darrow, d'Holbach, or Edwards, however. Unqualified endorsement of hard determinism has been rare. The principle at work seems to be that of the Victorian lady who, upon first hearing of Darwin's theory of evolution, exclaimed, "Descended from the apes. Let's hope it isn't true. But if it is, let's hope it does not become generally known."

Nonetheless, a core or kernel of the traditional hard determinist position persisted throughout the twentieth century and continues to play an important role in free will debates. To understand this kernel of hard determinism, note first that traditional hard determinism is defined by three theses: (1) free will is incompatible with determinism and (2) free will does not exist because (3) determinism is true. Modern thinkers who hold the kernel of hard determinism accept theses 1 and 2, but they are not committed to thesis 3—the universal truth of determinism. Aware of developments in twentieth-century physics, these modern thinkers are less confident than traditional hard determinists were that determinism is universally true in the natural world. They prefer to leave the question of the truth of determinism to the scientists. Yet they remain convinced that (1) free will and determinism are incompatible and that (2) free will (of the incompatibilist or libertarian kind) does not exist.

This is the kernel of traditional hard determinism—theses 1 and 2. What is interesting about this kernel is that it *amounts to a rejection of both compatibilism and libertarianism*. For anyone who accepts thesis 1 holds *against compatibilists* that free will is incompatible with determinism; and anyone who also accepts thesis 2 holds *against libertarians* that there is no free will of the true libertarian or incompatibilist kind. In short, those who hold this kernel of hard determinism are *skeptics* about free will. They reject both compatibilism and libertarianism, the traditional solutions to the free will problem. One such skeptic, Derk Pereboom, has introduced a useful expression to characterize those who accept theses 1 and 2. He calls them "hard incompatibilists."¹ They are "incompatibilists" by virtue of thesis 1 (true free will is not compatible with determinism) and "hard" by virtue of thesis 2 (true free will does not exist).

The skeptical positions of hard determinism and hard incompatibilism constitute a "third rail" in contemporary free will debates, the rail most people do not want to touch for fear of being electrocuted. For both these skeptical positions require living without belief in free will and true moral responsibility. Yet, while they may be unpopular, these skeptical positions are important because they pose a significant challenge to the other two main positions on free will, compatibilism and libertarianism.

3. Strawson's Basic Argument: The Impossibility of Moral Responsibility

But, you might ask: Why do modern skeptics about free will who are not committed to the truth of determinism believe that free will of the libertarian kind does not exist? In other words, why do they accept thesis 2 (free will does not exist) if they remain noncommittal about thesis 3 (that determinism is true)? The answer for most modern skeptics about free will is that they think free will in the libertarian sense is *impossible, whether determinism is true or not*. The most widely discussed skeptical argument to show this impossibility is an argument by Galen Strawson, which he calls the Basic Argument.² The idea behind Strawson's Basic Argument is an ancient idea: Having true free will of the libertarian kind would require that one be a *causa sui*—a cause of oneself. But being a *causa sui* is impossible, at least for us human beings. Strawson supports this idea with the following argument:

1. You do what you do because of the way you are (your nature or character).
2. To be truly responsible for what you do, you must be truly responsible for the way you are (for your nature or character).

3. But to be truly responsible for the way you are, you must have done something in the past for which you were also responsible to make yourself, at least in part, the way you are.
4. But if you were truly responsible for doing something in the past to make yourself what you are now, you must have been responsible for the way you were then (for your nature or character) at that earlier time.
5. But to have been responsible for the way you were at that earlier time, you must have done something for which you were responsible at a still earlier time to make yourself the way you were at that earlier time, and so on backward.

“Here one is setting off on a regress,” Strawson concludes, a regress that cannot go back forever in the case of human beings. Eventually you return to early childhood when your initial nature was not formed by you at all, but was the product of your heredity, early upbringing, and other factors beyond your control. Strawson then adds: “This argument goes through whether determinism is true or false. . . . Even if the property of being a *causa sui* is allowed to belong (entirely unintelligibly) to God, it cannot be plausibly supposed to be possessed by ordinary human beings.”³

Strawson then approvingly quotes Friedrich Nietzsche, who said:

The *causa sui* is the best self-contradiction that has been conceived so far; it is a sort of rape and perversion of logic. But the extravagant pride of man has managed to entangle itself . . . with just this nonsense. The desire for “freedom of the will” in the superlative metaphysical sense, which still holds sway, unfortunately, in the minds of the half-educated—the desire to bear the entire and ultimate responsibility for one’s actions oneself, and to absolve God, the world, ancestors, chance and society—involves nothing less than to be precisely this *causa sui* and, with more than Baron Munchausen’s audacity, to pull oneself up into existence by the hair, out of the swamps of nothingness.⁴

Baron Munchausen was the notorious teller of tales who claimed to have pulled himself from a ditch by his own hair. Needless to say, Nietzsche is another modern skeptic about free will who believes, along with Strawson, that the true free will of the ultimate libertarian kind is an illusion. Nietzsche thinks we should learn to accept our fate, even to learn to love our fate, and get on without the illusion of free will.

Is Strawson’s Basic Argument compelling? Premise 1 seems sound: “You do what you do because of the way you are (your nature or character).” As Hume pointed out, if our actions happened merely by accident or chance, if they did not flow from our character and motives, they could not

be imputed to us as “our” actions. How about premise 2? Is it the case that to be truly responsible for what you do, you must be truly responsible for the way you are (for your nature or character)? Think of McVeigh, Harris, and Klebold in connection with this premise. If we hold them responsible for their horrendous acts, it is because we think they were responsible, at least in part, for becoming the kinds of persons who would commit such acts. But this is what premise 2 requires—that McVeigh, Harris, and Klebold were at least in part responsible for becoming the kinds of persons who could commit such crimes. To hold them ultimately responsible we cannot think they were *entirely* shaped by psychological and social factors beyond their control.

Premise 3 seems sound as well: if McVeigh, Harris, and Klebold were responsible at least in part for being the way they were, it must have been because of something they *did* in the past for which they were responsible (some actions they performed or choices they made) to make themselves into the kinds of persons they became. But if premises 2 and 3 are sound, then steps 4 and 5 would seem to follow as well. For steps 4 and 5 simply reapply premises 2 and 3 to the past actions by which the agents made themselves what they are. If the agents are to be responsible for those past actions, they must also have been responsible for the characters and motives from which those past actions issued.

Is there any way to avoid Strawson’s conclusion from these plausible premises? It may be true, as his argument claims, that we cannot be creators of our “original” characters and motives—the characters and motives we began with in childhood before we ever made any free choices. But as we get older and develop, are we powerless to *change* the original characters we started with in childhood? Compatibilists and libertarians both respond to skeptical arguments like Strawson’s by saying that, although we are not the creators of our original characters, we can indeed freely change our natures and characters as we mature.

That seems like a piece of common sense. But Strawson replies that neither compatibilists nor libertarians give us an adequate account of *how* we could change our characters that accounts for true responsibility. If the *way* we change ourselves later in life, he argues, is *determined* by how *we already are*, as compatibilists allow, then that kind of change would not amount to true responsibility. But if the way we change ourselves later in life is *undetermined*, as libertarians require, then it would amount to mere luck or chance and that would not be true responsibility either. In other words, Strawson accepts the objections to *both* compatibilism and libertarianism that were considered in chapters 3 and 4. To answer his Basic Argument, compatibilists or libertarians must succeed in answering the objections against their views of these chapters; and in

doing so they must show that one or another of their views can account for true responsibility.

4. *Living Without Free Will: Crime and Punishment*

We will be returning in later chapters to both compatibilist and libertarian attempts to account for true responsibility and thereby answer Strawson's challenge. But suppose, for the sake of argument, that skeptical arguments against free will, such as Strawson's, cannot be answered. Can we live without the illusion of free will, as Nietzsche says we must? Skeptics about free will have addressed this question; and many of them have argued that living without the illusion of free will would not have the dire consequences that proponents of free will claim. Some skeptics about free will have gone even farther, affirming, as Nietzsche does, that giving up the illusion of free will would actually lead to a more positive, healthy, and honest approach to life.

Ted Honderich is one such skeptic who has addressed the consequences of living without free will.⁵ Honderich concedes that if we believed, as he does, that our behavior was sufficiently determined that we lacked free will, we would have to give up some important "life-hopes," but not all life-hopes. For example, we could no longer believe that our successes and accomplishments were really "up to us" in the sense that we were the ultimate "originators" of our actions. Nor could we believe that we were ultimately responsible for the traits of character in which we took pride—that we were hardworking, diligent, loyal, successful, and so on. To the extent that we had such characteristics, we would have to admit that we were merely lucky in our heredity and formative circumstances.

But most everyday life-hopes would remain, says Honderich. Desires to become a successful actor or dancer or writer, to start a business, to find love, to have children, to be admired by others—these hopes that give meaning to life would not be undermined by the belief that we are not the "originating" causes of our own characters. What these everyday life-hopes require is only that, if we make the appropriate voluntary efforts, there is a good chance that nothing will prevent us from realizing our cherished goals. Even if our behavior is determined, we cannot know in advance how things are destined to turn out. So we must go on trying to realize our life-hopes and dreams in the same manner as we would if we did believe we had free will in the incompatibilist sense, though in fact we do not.

How does this skeptical view of Honderich's differ from compatibilism? Honderich says that compatibilists try to convince us that if determinism

were true, nothing of importance would be lost in the way of freedom and responsibility. But this, Honderich thinks, is mistaken. Life-hopes that depend on believing that we are the undetermined originators of our characters and actions *are* important to our self-image. We are in fact giving up something important when we take a hard determinist or hard incompatibilist position. We should be honest and not deceive ourselves about that. But enough life-hopes remain, he thinks, to permit us to go on living in meaningful ways.

How would we deal with criminal behavior if we took this skeptical position on free will? According to Honderich, we would have to give up a *retribution* theory of punishment. According to the retribution theory, punishment of criminal behavior is right because it is *deserved*. The criminal has done wrong and must repay in kind for the wrong inflicted. "An eye for an eye" is the motto of the retribution theory. But if persons lacked free will, they would not be ultimately blameworthy for their actions and therefore punishment would not be truly deserved. So if hard determinism or hard incompatibilism were true, the retribution theory of punishment would have to be given up.

But Honderich insists that giving up the retribution theory does not mean we have to stop punishing criminals. There are other justifications for punishment that remain valid even if free will is rejected. The most common of these alternative justifications is *deterrence*. We also punish criminals to discourage them from committing future crimes and, even more important, we punish them to deter other persons from committing similar crimes. Still another motive for punishment is to *reform* or *rehabilitate* criminals so that they will return from prison as productive members of society. These motives for punishment—deterrence and reform—remain legitimate, Honderich insists, even if we reject free will. So we need not fear that our prisons would be emptied if everyone came to believe that people lack free will. Indeed, Honderich suggests that, if we gave up a belief in free will, we would put more emphasis on the prevention of crime through deterrence and reform rather than on retribution and vengeance—and society would be better off as a result.

Another skeptic about free will, Derk Pereboom, takes Honderich's arguments about criminal punishment a step further. In his book, aptly titled *Living Without Free Will*, Pereboom introduces a quarantine analogy to justify criminal punishment:

Ferdinand Schoeman has argued that, if in order to protect society, we have the right to quarantine people who are carriers of severe communicable diseases, then we also have the right to isolate the criminally dangerous to protect society. . . . This is true irrespective of the carriers' moral responsibility

for the disease. If a child is a carrier of the Ebola virus by virtue of its being passed on to her at birth from her parent, quarantine is nevertheless intuitively legitimate.

Furthermore, if we have the right to "quarantine" criminals, we have the right to tell people in advance that they will be isolated from society if they commit crimes. . . . This publicity itself has a powerful deterrent effect.⁶

An advantage of the quarantine model cited by Pereboom is that punishments would not be more severe than is needed to protect society and deter future crime, just as a quarantine of the sick should not be more restrictive than is needed to protect society from diseases. But a difficulty of the quarantine model is that it might allow us to jail persons who have not committed any crime but yet are thought to be a danger to society.

In response to this objection, Schoeman argues that it is more difficult to predict who will commit future crimes than it is to determine who has a dangerous communicable disease. But while this may usually be the case, is it always the case? There are some very bad and potentially dangerous people out there. (Consider the debates about how to treat child molesters who have been released from prison after serving time for their crimes.) Retributivists would argue, in response, that practices of punishment are bound to be unfair if we do not focus on who *deserves* to be punished, but instead focus only on what punishments will deter crime or protect society. If the focus is entirely on deterrence and protection rather than on retribution, injustices are bound to arise. Pereboom responds that the quarantine model works pretty well in most cases. If we reject free will, we would have to live with the few cases in which the quarantine model might be unfair. After all, those who are quarantined because they are sick are usually innocent as well. Also, if we place a high value on freedom, we will be reluctant for that reason alone to jail people who have not actually committed a crime.

5. *Personal Relations: Love, Admiration, and All That*

How would the rejection of free will affect our personal relations? Would the value of a person's love for you be deflated if you came to believe the person was determined to love you by heredity and environment? Many people think so because, as Pereboom says: "One might argue that we very much want to be loved by others as a result of their free will—we want freely willed love." But, he adds: "Against this, the love parents have for their children is typically engendered independently of the parents' will and we do not find this love deficient."⁷ Also, when we fall in love romantically,

it is rarely a matter of our free decision. Yet we do not find romantic love less satisfying for that reason. But is there not a mature kind of love we desire from lovers, spouses, friends, and even parents when we are older that would be deficient if we knew that factors beyond the others' control determined that they love us? To this objection, which I once posed to Pereboom's position, he responds as follows:

If we indeed desire a love of this kind, then we desire a kind of love that is impossible if hard incompatibilism is true. Still the kinds of love that are invulnerable to hard incompatibilism are surely sufficient for good relationships. If we aspire to the sort of love parents typically have toward their children, or the kind romantic lovers ideally have . . . or the type shared by friends . . . whose relationship is deepened by their interactions, then the possibility of fulfillment in personal relationships is far from undermined [by hard incompatibilism].⁸

Similar questions arise about other attitudes besides love. Could we admire people for generous or heroic deeds if we did not think they were ultimately responsible for those deeds? Could we feel grateful to them? Could we resent them or blame them if they reacted treacherously or deceitfully toward us? Pereboom says that some of these reactive attitudes (such as blame and guilt) would have to be given up if we accepted hard determinism or hard incompatibilism. But other significant attitudes of these kinds would not have to be given up altogether. We could go on believing that acts of certain kinds, say, of generosity and heroism, are admirable and that acts of other kinds are despicable even if we not believe that persons are ultimately responsible. Gratitude, for example, he says, "typically involves joy occasioned by the beneficent act of another. But hard incompatibilism fully harmonizes with being joyful and expressing joy when others are considerate and generous on one's behalf."⁹

6. *Illusion and Free Will*

Thus, Honderich and Pereboom believe we can live meaningful lives without the illusion of free will, though some important hopes and attitudes would have to be changed. But another skeptic about free will is not so confident that we can live meaningfully without belief in free will. Saul Smilansky agrees with Honderich and Pereboom that free will and determinism are incompatible and that libertarian free will does not exist. That is, he also holds theses 1 and 2 of section 2, the kernel of hard determinism. But Smilansky thinks Honderich and Pereboom are too optimistic

about the possibilities of living without belief in such a free will. So in his book *Free Will and Illusion*, Smilansky makes the provocative suggestion that even though we do not have true free will and moral responsibility in the deeper incompatibilist sense, we must foster the illusion in people that we do.¹⁰ He says:

To put it bluntly: People as a rule ought not to be fully aware of the ultimate inevitability of what they have done, for this will affect the way in which they hold themselves responsible. . . . We often want a person to blame himself, feel guilty and even see that he deserves to be punished. Such a person is not likely to do all this if he internalizes the ultimate hard determinist perspective, according to which . . . he could not strictly have done anything else except what he did do.¹¹

Smilansky wonders whether society as we know it could survive if most people came to believe that they were not truly responsible for their behavior. Some people might become more humane and understanding in their treatment of others knowing that no one was ultimately responsible. But Smilansky suggests that most people might simply become more selfish and no longer feel restrained by the requirements of morality. The stability of civilized societies would then be threatened. Only force and fear of punishment would keep people from breaking the law. As one of America's founders, James Madison, argues in Federalist Paper 10, if society has no ethical foundation, the law alone will not protect us. Smilansky also argues that accepting the hard determinist or hard incompatibilist perspective would be "extremely damaging to our view of ourselves, to our sense of achievement worth and self-respect."¹² Contrary to the arguments of Honderich and Pereboom, he thinks that giving up certain reactive attitudes such as blame, guilt, and resentment would have dire effects for society and personal life.

All this suggests to Smilansky that we must foster the illusion of free will and moral responsibility. (As the Victorian lady said of Darwin's theory: "If it is true, let us hope it does not become generally known.") Smilansky does not mean that we should induce illusory beliefs in the masses, in the manner of the movie *The Matrix* in which almost everyone lives in a virtual, computer-created, illusory world. Rather he thinks the illusion of free will is already in place. For most people already think of themselves either as compatibilists or libertarians. But compatibilists believe we already have all the freedom and responsibility we need even if determinism is true. And libertarians believe we also have the deeper incompatibilist free will. Both are wrong, according to Smilansky. But he thinks these illusory beliefs play a largely positive social and moral role and we should leave them in place rather than undermining them.

I will leave the reader to judge who wins this debate. Can we live meaningful lives without the illusion of free will and ultimate moral responsibility, as hard determinists or hard incompatibilists such as Honderich, Pereboom, Strawson, and Nietzsche say we must? Would the moral foundations of society survive intact? If not, could we really live in illusion, as Smilansky counsels us to do, if we knew the truth? What if people in *The Matrix* all found out it was all a dream?

Suggested Reading

Galen Strawson's Basic Argument against the intelligibility of free will appears in *Freedom and Belief* (Oxford, 1986) and in his 1994 essay "The Impossibility of Moral Responsibility," reprinted in Gary Watson's edited volume, *Free Will*, 2nd ed. (Oxford, 2003). Ted Honderich's view is most clearly presented in *How Free Are You?* (Oxford, 1993). Derk Pereboom's hard incompatibilist view is developed in his book *Living Without Free Will* (Cambridge, 2001), and Saul Smilansky's illusionist view is developed in his *Free Will and Illusion* (Oxford, 2000).



Free Will and Modern Science

1. Introduction

Can we make sense of a free will that requires Ultimate Responsibility of the kind described in the preceding chapter? Many philosophers think not. They argue (in the manner of Nietzsche and Strawson in chapter 7) that being the *ultimate* source of one's will and actions is an incoherent and impossible ideal, since it would require us to be "prime movers unmoved" or "uncaused causes of ourselves"—"the best self-contradiction that has been conceived so far," as Nietzsche put it. Ultimate Responsibility, or UR, requires that there be some acts in our lifetimes that do not have sufficient causes or motives. But how could acts having neither sufficient causes nor motives be free and responsible actions?

In chapter 5, I noted that traditional libertarian theories of free will have usually appealed to "extra factors" in response to these problems. Realizing that free will cannot merely be indeterminism or chance, libertarians have introduced additional and often mysterious forms of agency or causation to make up the difference, such as immaterial minds, noumenal selves outside space and time or non-event agent-causes. The idea behind such extra-factor strategies is easy enough to understand: since indeterminism leaves it open which way an agent will chose or act, some "extra" kind of causation or agency must be posited over and above the natural flow of events to account for the agent's going one way or the other—something else must tip the balance. This is a tempting way to think. But introducing extra forms of causation or agency beyond the natural flow of events has invited charges that libertarian theories of free will are obscure and mysterious and cannot be reconciled with modern scientific views about human beings.

Libertarians in general have not done a good job explaining how their view of free will can be reconciled with modern scientific views about human beings and the cosmos. This is the challenge I want to take up in the present chapter. Can a libertarian view of free will requiring Ultimate Responsibility be made intelligible without appealing to obscure or mysterious forms of agency or causation? Can such a free will be reconciled with what we know about human beings in the modern physical, biological, and human sciences? To answer these questions, I believe we have to rethink issues about freedom, responsibility, and indeterminism from the ground up, without relying on appeals to extra factors unless absolutely necessary. What follows is my own attempt to do this. Consider it a proposal meant to stimulate thinking about how free will might exist in the natural world where we humans exist and must exercise our freedom.

2. Physics, Chaos, and Complexity

We must grant, first of all, that if any libertarian theory of free will is to succeed there must be some genuine indeterminism in nature to make room for it. As the ancient Epicurean philosophers said, the atoms must sometimes "swerve" in undetermined ways if there is to be room in nature for free will. Moreover, it would be no use if the atoms swerved in outer space somewhere far from human affairs. They must swerve where it would matter for human choice and action, for example, in the brain. This is true even if one postulates special kinds of agent-causes or a non-material self to intervene in the brain. If these special forms of agency are to have any room to operate, the indeterminism must be there to begin with.

This is the point, as we have seen, where some scientists want to bring modern quantum physics into the picture to help account for free will. Suppose there were quantum jumps or other undetermined quantum events occurring in the brain. We know that information processing in the brain takes place through the firing of individual neurons or nerve cells in complex patterns. Individual firings of neurons in turn involve the transmission of chemical ions across neuronal cell walls, stimulated by various chemicals, called neurotransmitters, and by electrical stimuli coming from other neurons. Some neuroscientists have suggested that quantum indeterminacies in the transmission of these chemical ions across the cell walls of neurons might make the exact timing of the firings of individual neurons uncertain, thus introducing indeterminism into the activity of the brain and making "room" for free will.

Such suggestions are speculative. But even if they were correct, how would they help with free will? It was noted earlier that if choices were to occur as the result of quantum jumps or other undetermined events in the brain, the choices would not be under the control of the agents and would scarcely count as free and responsible actions. A similar criticism was made of the ancient Epicurean view. How could the chance swerve of atoms help to give us free will? Another problem about using quantum indeterminacy to defend free will was also mentioned in chapter 1. Determinists, such as Honderich, point out that quantum indeterminacy is usually insignificant in the behavior of larger physical systems like the human brain and body. When large numbers of particles are involved, as in the transmission of chemical ions across cell walls, any quantum indeterminacies would most likely be “damped” out and would have negligible effects on the larger activity of the brain and body.

Maybe so. But there is another possibility suggested by some scientists. Quantum theory alone will not account for free will, they concede. But perhaps quantum physics could be combined with the new sciences of “chaos” and “complexity” to help make sense of free will. In “chaotic” physical systems, very small changes in initial conditions lead to large and unpredictable changes in the system’s subsequent behavior.¹ You may have heard the narrative in which the fluttering of a butterfly’s wings in South America initiates a chain of events that affects weather patterns in North America. Perhaps that famous example is something of an exaggeration. But chaotic phenomena, in which small changes lead to large effects, are now known to be far more common in nature than previously believed, and they are particularly common in living things. There is growing evidence that chaos may play a role in the information processing of the brain, providing some of the flexibility that the nervous system needs to adapt creatively—rather than in predictable or rigid ways—to an ever-changing environment.

Determinists, to be sure, are quick to point out that chaotic behavior in physical systems, though unpredictable, is usually deterministic and does not itself imply genuine indeterminism in nature. But some scientists have suggested that a combination of chaos and quantum physics might provide the genuine indeterminism one needs. If the processing of the brain does “make chaos in order to make sense of the world” (as one recent research paper puts it²), then the resulting chaos might magnify quantum indeterminacies in the firings of individual neurons. These chaotically magnified indeterminacies in the firings of neurons would have large-scale indeterministic effects on the activity of neural networks in the brain as a whole. The indeterminacy at the neuron level would no longer be “damped out,” but would have significant effects on cognitive processing and deliberation.

But once again we might ask how even this would help with free will. If indeterminacy in our neurons were amplified to have significant effects on our mental processing and deliberation, would that give us any greater control and freedom? More likely it would give us less control and freedom. Wouldn’t deliberation become something like spinning a roulette wheel in one’s mind to make a choice? Maybe. But before we jump to conclusions, we need to look more deeply into the situation. If there were some significant indeterminism available in the brain, could we make more sense of it than simply spinning roulette wheels? Let us see. What is required to answer these questions, as I suggested, is a thorough rethinking of issues about freedom, responsibility, and indeterminism.

3. *Conflicts in the Will*

The first step in this rethinking 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 noted in chapter 11. Not all acts done of our own free wills have to be undetermined, only those acts by which we made ourselves into the kinds of persons we are—namely, the “will-setting” or “self-forming actions” (SFAs) that are required for ultimate responsibility.

Now I believe that 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 may be faced with difficult tasks for which we have aversions. In all such cases of difficult self-forming choices in our lives, we are faced with competing motivations and have to make an effort to overcome the temptation to do something else we also strongly want. There is tension and uncertainty in our minds about what to do at such times, let us suppose, that is reflected in appropriate regions of our brains by movement away from thermodynamic equilibrium—in short, a kind of “stirring up of chaos” in the brain that makes it sensitive to micro-indeterminacies at the neuronal level. The uncertainty and inner tension we feel at such soul-searching moments of self-formation would thus be reflected in the indeterminacy of our neural processes themselves. What we experience internally as uncertainty about what to do on such occasions would correspond physically to the opening of a window of opportunity that temporarily screens off complete determination by influences of the past.

When we do decide under such conditions of uncertainty the outcome is not determined, thanks to the indeterminacy that preceded it. Yet the

outcome can be willed either way we choose, rationally and voluntarily, because in such self-formation, the agents' prior wills are divided by conflicting motives. Consider a businesswoman who faces a conflict of this kind. She is on her way to an important meeting when she observes an assault taking place in an alley. An inner struggle arises between her conscience on the one hand (to stop and call for help for the assault victim) and her career ambitions, on the other hand, which tell her she cannot miss this important business meeting. She has to make an effort of will to overcome the temptation to do the selfish thing and go on to the meeting. If she overcomes this temptation, it will be the result of her effort to do the moral thing; but if she fails, it will be because she did not *allow* her effort to succeed. For while she willed to overcome temptation, she also willed to fail. That is to say, she had strong reasons to will the moral thing, but she also had strong reasons, ambitious reasons, to make the selfish choice that were different from, and incommensurable with, her moral 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*. Thus the choice we eventually make, though undetermined, can still be rational (made for reasons) and voluntary (made in accordance with our wills), whichever way we choose.

Now let us add a further piece to the puzzle. Just as indeterminism need not undermine the rationality and voluntariness of choices, so indeterminism in and of itself need not undermine control and responsibility. Suppose you are trying to think through a tough math problem. Say there is an indeterminacy in your neural processes complicating the task. This indeterminacy would make your task more difficult, in much the same way that low background noise would be slightly distracting if you were trying to solve a tough math problem. Whether you are going to succeed in solving the problem is uncertain and undetermined because of the distracting neural noise. Yet, if you manage to concentrate and solve the problem nonetheless, we have 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 many examples supporting this idea of indeterminism functioning as an obstacle to success without precluding responsibility. Included among these examples are the Austin-style examples discussed in chapter 11. Recall the assassin, who was trying to shoot the prime minister but might miss because undetermined events in his nervous system might lead to a jerking or wavering of his arm. If the assassin did 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. The indeterminism here functioned as an obstacle to his success but did not rule out his responsibility *if* he succeeded.

Here is another example. A husband, beside himself with rage while arguing with his wife, swings his arm down on her favorite glass-top table, 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 undetermined whether the table will break right up to the moment it is struck. Whether the husband breaks the table is undetermined, and yet he is clearly responsible if he does break it. (It would be a poor excuse to offer his wife if he claimed, "Chance did it, not me." Though indeterminism was involved, chance didn't do it, *he* did.) In this example as in the preceding one, the agent can be held responsible for an action even though the action is undetermined.

Now these examples—of the math problem, the assassin, and the husband—are not all we want for free will. They do not amount to genuine exercises of self-forming actions (SFAs) like the businesswoman whose will is divided between conflicting motives. The businesswoman wants to help the assault victim, but she also wants to go on to her meeting. By contrast, the assassin's will is not equally divided. He wants to kill the prime minister, but he does *not* also want to fail. (If he fails therefore, it will be *merely* by chance.) So while the examples of the assassin, the husband, and the like do not tell us all we need to know about free will, they do provide some clues to what free will requires. To go further, we have to appeal to some additional ideas.

4. *Parallel Processing*

Imagine in cases of conflict characteristic of self-forming actions or SFAs, like the businesswoman's that the indeterministic noise, which is providing an obstacle to her overcoming temptation, is coming not from an external source but 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 the woman's conflicting motivations. (These neural networks are complex networks of interconnected neurons in the brain, circulating impulses in feedback loops that are generally involved in higher-level cognitive processing.³) The input of one of these neural networks consists of the woman's reasons for acting morally and stopping to help the victim; the input of the other

network comprises her ambitious motives for going on to the meeting. The two neural networks are connected, so that the indeterministic noise, which is an obstacle to the woman's making one of her choices, is coming from her own desire to make the opposite choice. In these circumstances, when either of the pathways "wins" (i.e., reaches an activation threshold, which amounts to choice), the woman will be making her choice in spite of the indeterministic noise she had to overcome. Her choosing in spite of the noise obstacle will be like your solving the tough math problem in spite of distracting background noise. And just as we can say, when you solved the math problem by overcoming the distracting noise, that you did it and are responsible for it, so we can say this as well, I would argue, in the woman's case, *whichever way she chooses*. 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 pathway.

Note that under such conditions of indeterminism arising from conflicting alternatives, choices going either way will not be "inadvertent," "accidental," "capricious," or "merely random" (as critics of indeterminism say). On the contrary, the choices 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 that something is done "on purpose" rather than accidentally, capriciously, or merely by chance. Moreover, these conditions for saying the actions were done on purpose, taken together, I would argue, rule out each of the reasons we have for saying that agents act but do not have *control* over their actions. The agents need not have been acting under compulsion, coercion, constraint, inadvertence, accident, control by others, and so on.⁴ To be sure, we must grant that when choices are undetermined SFAs, agents do not control or determine which choice-outcome will occur *before* it occurs. But it does not follow that, because one does not control or determine which of a set of outcomes is going to occur before it occurs, one does not control or determine which of them occurs, *when* it occurs.

When the preceding conditions for SFAs are satisfied, and the agents exercise control over their future lives *then and there* by deciding, they have what I call *plural voluntary control* over the options in the following sense: the agents 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 in willing to do so, or otherwise controlled in doing or in willing to do so by any other agents or mechanisms. Each of these conditions can be satisfied for SFAs, like the businesswoman's, as I have described them. The conditions can be summed up by saying that the

agents can choose either way *at will*. In other words, the choices are "will-setting": we set our wills one way or the other in the *act* of deciding, and not before.

Note also that this account of self-forming choices amounts to a kind of "doubling" of the difficulty seen in the math problem example, where the agent had to make an effort to overcome indeterministic background noise. It is as if an agent faced with a self-forming 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 the businesswoman is trying to make a moral choice and to make a conflicting self-interested choice. The two competing choices correspond to two competing neural networks in her brain. 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 math problem, the agents will have succeeded in doing what they were knowingly and 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 might fail to break the table. But if these two agents *succeed*, despite the probability of failure, they are responsible, since they will have succeeded in doing what they were trying to do. And so it is, I suggest, with self-forming choices like the businesswoman's. The agents will be responsible *whichever way they choose* because whichever way they choose they will have succeeded in doing what they were trying to do. Their failure to do one thing is not a *mere* failure, but a voluntary success in doing the other.

Does it make sense to talk about an agent's trying to do two competing things at once in this way, or to solve two cognitive problems at once? Well, we now know that the brain is a "parallel processor"; it can simultaneously process 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 do something to help the victim is the same ambitious woman who wants to go to her meeting and make a sale. 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 outcome as *her* resolution of the conflict in her will, voluntarily and intentionally, not by accident or mistake.

5. Challenges to This View: Responsibility, Luck, and Chance

Obviously, many questions arise about the preceding view and a number of objections may be made to it. We cannot address all these questions and objections here, but let us consider some of the more important ones. Some people have objected that if choices like the businesswoman's really are undetermined, they *must* happen merely by chance—and so must be “random,” “capricious,” “uncontrolled,” “irrational,” and all the other things usually charged. The first step in responding to this objection is to question the assumption that if indeterminism is involved in an occurrence, that occurrence must happen *merely* as a matter of chance or luck. “Chance” and “luck” are terms of ordinary language that carry the meaning of “its being out of my control.” So using them already begs certain questions. “Indeterminism,” by contrast, is a technical term that merely rules out *deterministic* causation, but not causation altogether. Indeterminism is consistent with nondeterministic or probabilistic causation, where the outcome is not inevitable. It is therefore a mistake (in fact, one of the most common mistakes in debates about free will) to assume that “undetermined” means “uncaused” or “*merely* a matter of chance.”

A second objection is related to the first. One might argue that in the case of the businesswoman, since the outcome of her effort (the choice) is undetermined up to the last minute, she must have first made the 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 a mistaken image. On the view just presented, one cannot separate the indeterminism from the effort of will, so that *first* the woman's effort occurs, to be *followed* by chance or luck. One must think of the effort and the indeterminism as fused; the effort *is* indeterminate and the indeterminism is a property of the effort, not something separate that occurs after or before the effort. The fact that the effort has this property of being indeterminate does not make it any less the woman's *effort*. The complex recurrent neural network that realizes the effort in the brain is circulating impulses in feedback loops, and there is some indeterminacy in these circulating impulses. But the whole process is the woman's 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.” The woman 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.)

A third objection has to do with the notion of luck. If the businesswoman's efforts were undetermined, so that either effort might have failed, some critics argue, then it was just a matter of luck which effort succeeded. To address this by-now familiar objection, we need to look more closely at the issue of luck. Recall that one might say of the assassin and the husband that “they got lucky” in killing the prime minister and breaking the table, because their actions were undetermined and might have failed. Yet the surprising thing is that we still say the assassin and the husband were *responsible* if they succeeded in killing the prime minister or breaking the table. So we should ask ourselves the following question: why is it wrong to say “He got lucky, *so he was not responsible*” in the cases of the husband and the assassin? For it *is* wrong to say this, since they did get lucky and yet they were *still* responsible. (Imagine the assassin's lawyer arguing in the courtroom that his client is not guilty because his killing the prime minister was undetermined and might therefore have failed by chance. Would such a defense succeed?)

The first part of an answer to why the assassin and the husband are still responsible has to do with the point made earlier about “luck” and “chance.” These two words have question-begging implications in ordinary language that are not necessarily implications of “indeterminism” (for indeterminism implies only the absence of deterministic causation). The core meaning of “he got lucky” in the assassin and husband cases is “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 other meanings in these cases, meanings that are often associated in ordinary usage with “luck” and “chance,” the inference that a person “got lucky so he was not responsible” would not fail, as it clearly does. For example, if “luck” in these cases meant that the outcome was not his doing, or had occurred by mere chance, or that he was not responsible, then the inference “he got lucky so he was not responsible” would hold for the husband and assassin. But the point is that these further meanings of “luck” and “chance” do not follow *from the mere presence of indeterminism*.

The second reason why the inference “he got lucky, so he was not responsible” does not work in the cases of 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 were trying and wanting to do all along, 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), then (1) she will have "succeeded *despite the probability or chance of failure*," (2) 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 (3) when she succeeded (in choosing to return to help) her reaction was not "Oh dear, I did that by mistake, it was an accident; it was 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 the choice 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.

6. Choice and Agency

Here is a fourth objection that may have occurred to you. Perhaps we are begging the question by assuming that the outcomes of the woman's efforts are *choices* to begin with. If indeterminism is involved in a process (such as the woman's deliberation) so that its outcome is undetermined, one might argue that the outcome must merely *happen* and therefore cannot be somebody's *choice*. But there is no reason to assume that such a claim is true. A choice is the formation of an intention or purpose to do something. It resolves uncertainty and indecision in the mind about what to do. Nothing in such a description implies that there could not be some indeterminism in the deliberation and neural processes of an agent's preceding choice corresponding to the agent's prior uncertainty about what to do. Recall from our earlier arguments that the presence of indeterminism does not mean the outcome happened *merely* by chance and *not* by the agent's effort. Self-forming choices are undetermined but not uncaused. They are caused by the agent's efforts.

Well, say some critics, perhaps indeterminism does not undermine the idea that something is a *choice*, but rather that it is *the agent's* choice. This objection raises some important questions about agency. What makes the woman's choice her own on the foregoing account is that it results from her efforts and deliberation, which in turn are causally influenced by her reasons and her intentions (e.g., her intention to resolve indecision in one way or another). And what makes these efforts, deliberations, reasons, and intentions *hers* is that they are embedded in a larger motivational system realized in her brain in terms of which she defines herself as a practical

reasoner and actor. A choice is the agent's when it is produced intentionally by efforts, by deliberations, and by reasons that are part of this self-defining motivational system and when, in addition, the agent *endorses* the new intention or purpose, created by the choice, into that motivational system, making it a further purpose that will guide *future* practical reasoning and action.

Well then, say other critics, perhaps the issue is not whether an undetermined SFA, such as the businesswoman's, is a *choice*, or even whether it is the *agent's* choice, but rather how much *control* she has over it. For while it may be true, as argued earlier (in the discussion of plural voluntary control), that the presence of indeterminism need not eliminate control altogether, wouldn't it be the case that the presence of indeterminism at least *diminishes* the control persons have over their choices and actions? Is it not the case that the assassin's control over whether the prime minister is killed (his ability to carry out his purposes and do what he is trying to do) is lessened by the undetermined impulses in his arm? This criticism is related to a problem about libertarian freedom encountered in chapter 4. The problem is that indeterminism, wherever it occurs, seems to be a *hindrance* or *obstacle* to our realizing our purposes and hence is an obstacle to our freedom rather than an *enhancement* of it.

There is some truth to this objection. But I think what is true in it may reveal something important about free will. Perhaps we should concede that indeterminism, wherever it occurs, *does* diminish control over what we are trying to do and *is* a hindrance or obstacle to the realization of our purposes. But recall that in the case of the businesswoman (and SFAs generally), the indeterminism that is admittedly diminishing the agent's control over one thing she is trying to do *is coming from her own will*—from her desire and effort to do a different thing that she also wants to do. And the indeterminism that is diminishing her control over that different thing (in this case the selfish thing) is coming from her desire and effort to do its opposite (to be a moral person who acts on moral reasons). So, in each case, the indeterminism *is* in fact 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—the woman would indeed in a sense have "complete control" over one of her options. There would be no competing motives to stand in the way of her choosing it. But then also she would not be free to rationally and voluntarily choose the other option 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 for us to overcome.

Another objection to the preceding theory is that we are not consciously aware of making two competing efforts when we engage in self-forming choices. But the theory does not require that we be consciously aware of these competing efforts. The idea was to compare exercises of free will to other cases of parallel processing in the brain, such as vision. Neuroscientists tell us that when we see a visual object, such as a red barn, the brain actually processes different properties of the object (like shape and color) separately, through parallel pathways whose results are eventually brought together in the visual image. We are not introspectively aware of processing the redness of the barn and its shape separately and in parallel. In fact, this information about parallel processing in the brain comes as a surprise to us. But if these neurological theories are correct, that is what we are doing.

The preceding account of free will is suggesting that something similar may be going on when we make self-forming choices. We are not introspectively aware that our efforts (our efforts to make one or another of our competing choices succeed) are being processed on separate, though interacting, pathways in the brain; but that process may in fact be what is going on. If we actually introspected all that was going on when we made free choices, free will would be less mysterious and the problem of free will would be a lot easier to solve than it is. To solve it, we have to consider what may be going on behind the scenes when we are conscious of trying to decide about which of two options to choose and either choice is a difficult one because there are resistant motives pulling us in different directions.⁵

Let us conclude with one final objection to the account of free will presented in this chapter. This objection is perhaps the most telling and has not yet been discussed. It goes like this: even if one grants 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 considerable truth to this objection as well, but again I think it may be 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 in the garden of forking paths 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 a 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."

To sum up, in this chapter I have suggested how a libertarian free will requiring ultimate responsibility and indeterminism might be reconciled with current scientific knowledge. There is much to debate about the theory of this chapter and many objections can and have been made to it.⁶ I have tried to answer some of these objections here; but many other objections that also deserve answers have not been addressed. (Those who wish to pursue the issues further can look at the suggested readings that follow.) Many persons believe libertarian free will can never be reconciled with science and cannot exist in the natural order. Perhaps they will turn out to be right. But we should not conclude too hastily that free will of the deeper kind that libertarians believe in cannot be reconciled with science without first trying our best to see how it might be done.

Suggested Reading

The view presented in this chapter is further developed in my *The Significance of Free Will* (Oxford, 1996). Objections to this theory and further debates about it can be found in the references cited in note 6. An interesting collection of essays relating current research in the neurosciences, psychology and physics to free will is *The Volitional Brain*, edited by Benjamin Libet, Anthony Freeman, and Keith Sutherland (Imprint Academic, 1999). Other different attempts to reconcile free will with modern science include (from a libertarian perspective) David Hodgson "Hume's Mistake" (in *The Volitional Brain*, pp. 201–24) and Storrs McCall, *A Model of the Universe* (Oxford: Clarendon, 1994) and (from a compatibilist perspective) Henrik Walter, *Neurophilosophy of Free Will* (MIT, 2001), and Daniel Dennett, *Freedom Evolves* (Vintage, 2003).



Predestination, Divine Foreknowledge, and Free Will

1. Religious Belief and Free Will

Debates about free will are impacted by religion as well as by science, as noted in chapter 1. Indeed, for many people, religion is the context in which questions about free will first arise. The following personal statement by philosopher William Rowe nicely expresses the experiences of many religious believers who first confront the problem of free will:

As a seventeen year old convert to a quite orthodox branch of Protestantism, the first theological problem to concern me was the question of Divine Predestination and Human Freedom. Somewhere I read the following line from the Westminster Confession: "God from all eternity did . . . freely and unchangeably ordain whatsoever comes to pass." In many ways I was attracted to this idea. It seemed to express the majesty and power of God over all that he had created. It also led me to take an optimistic view of events in my own life and the lives of others, events which struck me as bad or unfortunate. For I now viewed them as planned by God before the creation of the world—thus they must serve some good purpose unknown to me. My own conversion, I reasoned, must also have been ordained to happen, just as the failure of others to be converted must have been similarly ordained. But at this point in my reflections, I hit upon a difficulty, a difficulty that made me think harder than I ever had before in my life. For I also believed that I had chosen God out of my own free will, that each of us is responsible for choosing or rejecting God's way. But how could I be responsible for a choice which, from eternity, God had ordained I would make at that particular moment of my life? How can it be that those who reject God's way do so of their own free will, if God, from eternity, destined them to reject his way?¹