

Chapter 13

Free Will

Overview

Determinism is the thesis that for any given state of the universe at a time there is exactly one possible resultant state that comes about because of the laws of nature. The problem of free will and determinism is the problem of reconciling either the truth or the falsity of determinism with the existence of free will and moral responsibility. The problem is based on the following claims:

- 1 Determinism is true, or determinism is false.
- 2 If determinism is true, then there is no free will.
- 3 If determinism is false, then there is no free will.
- 4 There is moral responsibility only if there is free will.
- 5 There is moral responsibility.

Claims (1)–(3) imply that there is no free will, but claims (4) and (5) imply that there is free will. Claims (1)–(5) are thus jointly inconsistent; they cannot all be true. There are nevertheless good reasons for endorsing each. The task for philosophers is thus to explain which of the claims is false and why the reasons in favor of it are not decisive.

Compatibilist theories try to reconcile free will and moral responsibility with the truth of determinism. Classic compatibilist theories, for instance, accept claims (4) and (5), but reject Claim (2). According to them, free will requires the ability to do otherwise, but the ability to do otherwise can be understood in a way that is compatible with determinism. Contemporary compatibilist theories, on the other hand, reject the idea that free will requires the ability to do otherwise.

Hierarchical theories, for instance, claim that free will can be given an account in terms of a hierarchy of higher-order desires, something that does not require the ability to do otherwise. Capacity-based theories, on the other hand, claim that free will can be given an account in terms of the exercise of certain capacities such as capacities for rational self-governance. Reactive attitude theories claim that moral evaluation depends on social practices. Since those practices do not depend on the truth or falsity of determinism, we can hold people morally accountable for their behavior regardless of whether or not determinism is true. Semicompatibilist theories also focus on the notion of moral responsibility; unlike most compatibilist theories, however, semicompatibilism accepts the incompatibility of free will and determinism. Free will requires the ability to do otherwise, semicompatibilists say, but moral responsibility does not. We can thus be morally responsible even if we do not have the ability to do otherwise – even if determinism is true. Semicompatibilists thus reject Claim (4), the claim that free will is necessary for moral responsibility.

By contrast with compatibilist theories, libertarian theories try to reconcile free will and moral responsibility with the falsity of determinism; they reject Claim (3). Simple indeterminist theories claim that free actions are completely uncaused events, but that we can still be held morally accountable for them because moral responsibility does not require that we actually bring about the actions we are responsible for. Causal indeterminist theories, on the other hand, deny that free actions are entirely uncaused. Free actions are caused by antecedent events, they say, but they are caused indeterministically. Antecedent events do not determine results; they merely make a certain result more probable. In addition, like causal indeterminist theories, agent causal theories claim that free actions have causes, but these causes are not antecedent events; they are instead agents themselves. There is, in other words, a special kind of causal relation – agent causation – that differs from event causation and that is responsible for bringing about actions. Finally, some libertarian theories affirm the existence of free will and moral responsibility, but deny the possibility of providing an informative positive account of how free will and moral responsibility are possible. The eighteenth-century philosopher Immanuel Kant argued for a position along these lines.

In addition to the foregoing theories, there are also some that deny the existence of free will and moral responsibility. Hard determinism endorses Claim (2) together with determinism. Free will and determinism are incompatible, it says, and determinism is true; consequently, free will does not exist. Hard incompatibilists, on the other hand, believe that the truth or falsity of determinism is an empirical matter – something it is up to scientists not philosophers to determine. Like hard determinists, they endorse Claim (2), and like hard determinists they deny that we are free and morally responsible, but their denial of free will and moral responsibility is based not on determinism, but on independent grounds.

Philosophers who deny free will and moral responsibility disagree about the practical implications of this denial. Some think denying the existence of free will

and moral responsibility requires significant changes in our social institutions and practices; others think it requires few changes if any.

13.1 The Problem of Free Will and Determinism

This book has focused so far on mind-body problems. In this chapter we consider another philosophical problem, the **problem of free will and determinism**. We take it for granted in our day-to-day dealings that we are free beings, that we are for the most part in control of our actions, and hence that we are largely responsible for what we do. We make this assumption whenever we speak of choices, or hold people accountable for the actions they choose to perform. The assumption that we choose to do things is the reason we feel justified in imprisoning people or in forcing them to pay fines or simply blaming them for their behavior. It is also the reason we feel justified in praising people for their actions, or giving them credit for doing the right thing in difficult circumstances. If what people did were not a matter of choice, if it were not up to them, then we could not hold them accountable for their behavior; we would not be justified in giving them either moral praise or moral blame. Moral accountability is based on the assumption that there is such a thing as free will or choice.

It is nevertheless difficult to understand how free will or choice can exist if the universe is the type of place modern science reveals it to be. Our best science tells us that the universe contains matter and energy, that there are atoms, molecules, quarks, leptons, photons, and other physical particles, and that these particles are distributed throughout the universe in various ways. Some constitute this computer, others constitute this chair or the Earth, and yet others constitute us – people. Science tells us that the states of matter and energy undergo changes with time. Water, for instance, changes from a liquid state to a solid state if it gets cold enough, and it changes from a liquid state to a gaseous state if it gets hot enough. Likewise, the energy stored in the chemical bonds of a quantity of gasoline is converted into heat in an automobile engine, which is transformed in turn into mechanical energy by the pistons, and back again into heat when the brakes bring the automobile to a halt. Science informs us, moreover, that these changes happen in very regular, very predictable ways. It also informs us that they do not happen by choice. The physical particles that constitute things do not choose to behave as they do; their behavior is not free. It is instead governed completely by the laws of physics, and these laws have nothing to do with freedom or choice. This seems to be the type of world our best science reveals – a world in which the ultimate constituents of things, the basic particles out of which everything is constituted, are not free beings who make choices.

A philosophical problem arises when we try to reconcile this view of the universe with human freedom, for if the universe is the kind of place science reveals

it to be, then it is difficult to see how human behavior could be governed by free will or choice. Consider a human action such as Madeleine's breaking the dishes out of anger. We can hold Madeleine accountable for breaking the dishes only if we believe she chose to break the dishes – only if we believe she had the capacity to act in different ways, and exercised that capacity by acting one way as opposed to another. Clearly, however, Madeleine's action consisted in the movements of her limbs – the movements of her arms as she cast the dishes to the floor, say. But those arm movements consisted in nothing but the movements of various physical particles – the particles composing her muscles, bones, skin, and nerves. None of these particles acted freely; none of them changed its position on account of a decision or choice to move one way instead of another. But if Madeleine's action consisted in nothing but changes in the positions of particles, and none of these changes happened freely, it is difficult to see how Madeleine's action as a whole could have happened freely. And if it is difficult to see how Madeleine's action could have happened freely, it is difficult to see how we could hold Madeleine accountable for it.

The problem that arises here centers on the compatibility of free will and **determinism**. Determinism is the thesis that for any given state of the universe at a time there is exactly one possible resultant state that comes about as a result of the laws of nature. In the problem as we've described it, the total state of the universe at a time consists in the distribution of all matter and energy in the universe at that time, and the laws of nature are the laws of physics. According to determinism, the laws of nature completely fix or determine what will happen in the universe from one moment to the next. A being who knew all the laws of nature and the total distribution of matter and energy in the universe at a given time would be able to predict with perfect accuracy any subsequent state of the universe. A universe exactly like ours with the same laws of nature and the same initial conditions would be guaranteed to have the same history as our universe. If determinism is true, the same past states guarantee the same future ones.

If determinism is true, then it appears that free will cannot exist. Consider a human action such as Madeleine's breaking the dishes out of anger. We can hold Madeleine accountable for breaking the dishes only if we believe she was capable of doing otherwise – only if we think she didn't *have* to break the dishes, but instead *chose* to break the dishes. Her anger, however, was included in the state of the universe at the time she broke the dishes, and determinism claims that for any given state of the universe at a time there is only one possible resultant state. If determinism is true, then given all the conditions of the universe at the time, including Madeleine's anger, there was only one possible state that could have resulted, and that resultant state included Madeleine's breaking the dishes. Under the circumstances, then, there was only one thing Madeleine could have done: break the dishes. Neither her action, nor anything else that happened in the universe at that time could have turned out otherwise than the way it turned out in fact. Madeleine thus *had* to break the dishes – she was, as it were, forced to do so

by the laws of nature. If determinism is true, therefore, it appears that there can be no such thing as free will, for free will requires the possibility that things might have turned out otherwise; it requires, for instance, that Madeleine could have done something other than break the dishes, but this is a possibility determinism rules out.

Someone who is discouraged by the conclusion that there is no such thing as free will, and hence no such thing as moral responsibility, might try to avoid this result by rejecting determinism: "If determinism rules out free will," someone might say, "that just means that determinism must be false! The laws of nature do not force things to be one way instead of another. Problem solved!" Unfortunately, this attempt at solving the problem generates a new problem. Recall that holding people accountable for their actions requires that those actions be in their control; it requires that they actually be responsible for bringing about those actions. If determinism is false, however, it starts to look like people are never responsible for bringing about their actions. Why is that the case?

Consider what it really means for determinism to be false. If determinism is false, then for any given state of the universe at a time, there is more than one possible resultant state. Suppose, for instance, that at time t_1 , the total state of the universe includes Madeleine's being angry, and at time t_2 the total state of the universe could either include Madeleine's breaking the dishes or Madeleine's not breaking the dishes. Suppose, finally, that Madeleine breaks the dishes in fact. We can now ask: Why did Madeleine break the dishes at t_2 instead of not breaking the dishes at t_2 ? What was responsible for the universe turning out one way (with the broken dishes) instead of the other? Initially we might want to say that this came about because Madeleine *chose* to break the dishes. In other words, Madeleine's choice brought it about that the universe ended up with the broken dishes as opposed to having ended up without them. But to say that Madeleine's choice is what brought about her dish-breaking – to say, in other words, that her choice is what made the universe turn out one way as opposed to another – seems like another way of saying that her choice *determined* a certain result. When we explain how the dishes ended up on the floor in a broken state at time t_2 by appeal to Madeleine's choice, we seem to be appealing to a determining factor – a choice – which brings about a certain outcome. Madeleine's choice rules out the alternative state of affairs; it determines the breaking of the dishes. So if we appeal to choice to explain the breaking of the dishes, it seems that we remain committed to a deterministic picture of things.

If we *really* want to reject determinism, we need to reject the idea that there is a factor such as choice that explains why the universe ended up one way at t_2 as opposed to another. In order for determinism to be false, we need to suppose that *nothing* explains why Madeleine ends up breaking the dishes at t_2 as opposed to not breaking the dishes at t_2 . We must be committed to saying, in other words, that Madeleine's breaking the dishes just *happened*. She didn't bring about that result; if determinism is false, nothing brought about that result. If determinism

is false, nothing brings about any result; things just happen as they do without the influence of any determining factors. Notice now what consequences this has for human action: if nothing is responsible for bringing about what happens, then we are not responsible for bringing about our actions. Madeleine, for instance, is not responsible for breaking the dishes. If she didn't bring about that result, if that result was not something that was in her control, then she cannot be held accountable for that result any more than she can be held accountable for the weather, for in neither case is she responsible for bringing it about.

If determinism is false, therefore, nothing brings about one possible state of the universe as opposed to another. We are not in control of our actions; we do not bring them about; they just happen. But if our actions just happen, if we do nothing to bring them about, if they are not in our control, then we cannot be held accountable for them since moral responsibility requires that we bring about one thing as opposed to another. If determinism is false, therefore, there can be no moral responsibility. One philosophical problem has thus taken the place of another: in place of a problem reconciling free will with the truth of determinism, we face a problem reconciling free will with the falsity of determinism.

If the universe is the way science reveals it to be, therefore, if it is a sea of matter and energy governed by physical laws, then it is difficult to see how a phenomenon such as free will or choice can exist, and if it is difficult to see how a phenomenon such as free will or choice can exist, then it is difficult to see how we can be justified in holding people accountable for their behavior. Notice, moreover, that the problem cuts even deeper than the foregoing considerations suggest, for we would face the very same problem even if we were not physical beings – even if, say, a view like substance dualism (Sections 3.1–3.2) or idealism (Sections 9.1–9.2) were true. To appreciate this, we need only appreciate that determinism does not specify what the total state of the universe at a time consists in or what exactly the laws of nature are. In describing the problem of free will and determinism so far, we have taken the total state of the universe at a time to consist in the distribution of all matter and energy in the universe at that time, and we have taken the laws of nature to be the laws of physics. But we need not understand the universe and the laws of nature in this way.

Imagine, for instance, that substance dualism is true. Imagine the universe consists of not only physical things but also nonphysical ones, and the laws of nature include not just the laws governing physical things but also laws governing nonphysical ones. In that case, we face the same problem explaining the possibility of free will and moral responsibility, for the laws governing physical and nonphysical things will either be deterministic or not. If they are deterministic, then they will rule out free will and moral responsibility for the reasons described earlier: free will and moral responsibility require the ability to do otherwise, but determinism rules that out. If, on the other hand, they are not deterministic, they will rule out free will and moral responsibility because free will and moral responsibility require control, and without deterministic causal relations, agents are not in

control of their actions. The problem of free will and determinism does not concern only physical things, therefore; it concerns the universe and the laws of nature in general whatever the universe and laws of nature happen to be.

The problem of free will and determinism can be formulated as a set of five jointly inconsistent claims:

- | | | | | |
|---|---|---------|---|------------------------|
| 1 Determinism is true, or determinism is false. | } | implies | → | There is no free will. |
| 2 If determinism is true, then there is no free will. | | | | |
| 3 If determinism is false, then there is no free will. | | | | |
| 4 There is moral responsibility only if there is free will. | } | implies | → | There is free will. |
| 5 There is moral responsibility. | | | | |

Claims (1)–(3) imply that there is no free will, but claims (4) and (5) imply that there is free will. Claims (1)–(5) are thus jointly inconsistent; they cannot all be true. There are nevertheless good reasons for endorsing each. The task for philosophers is thus to explain which of these claims is false, and why the reasons in favor of it are not decisive. Let us briefly survey some of the standard solutions to the problem.

13.2 Some Solutions to the Problem

Solutions to the problem of free will and determinism look to reject one or more of the problem's claims (1)–(5). **Compatibilist theories** try to reconcile free will and moral responsibility with determinism. *Classic compatibilist theories*, for instance, reject Claim (2) (Section 13.3). They accept that there is moral responsibility and that moral responsibility requires free will, and they thus accept that free will exists, but they deny that the existence of free will is incompatible with determinism. Free will requires the ability to do otherwise, they say, but that ability can be understood in a way that is compatible with determinism.

It is important to note that compatibilism is not committed to determinism. Compatibilism is a claim about whether the existence of free will is compatible with the truth of determinism. It does not imply that determinism is true. In fact, many compatibilists do not take a stand on the truth or falsity of determinism. They claim that the truth or falsity of determinism is an empirical matter not a philosophical one – that it is up to scientists not philosophers to determine whether or not determinism is true in fact. Compatibilism is thus neutral with regard to the truth or falsity of determinism. Some compatibilists nevertheless do endorse determinism. Their view is often called **soft determinism**; it conjoins

compatibilism with determinism. Every soft determinist is a compatibilist, therefore, but not every compatibilist is a soft determinist.

Contemporary compatibilist theories try to reconcile moral responsibility with determinism as well, but unlike classic compatibilist theories, they reject the idea that free will requires the ability to do otherwise. **Hierarchical theories**, for instance, claim that free will can be understood in terms of a hierarchy of higher-order desires, something that does not require the ability to do otherwise (Section 13.5). **Capacity-based theories**, on the other hand, claim that free will can be understood in terms of the exercise of certain capacities such as the capacity for rational self-governance (Section 13.5), and this too does not require the ability to do otherwise. **Reactive attitude theories**, for their part, focus on the notion of moral responsibility (Section 13.6). Moral evaluation, they say, depends on social practices that we cannot help engaging in because those practices satisfy basic human needs. Since those practices do not depend on the truth or falsity of determinism, we can hold people morally accountable for their behavior regardless of whether or not determinism is true. **Semicompatibilist theories** focus on the notion of moral responsibility as well (Section 13.6). Unlike most compatibilists, however, semicompatibilists accept Claim (2); they claim that free will is incompatible with determinism. They instead reject Claim (4); they deny that free will is necessary for moral responsibility. Free will requires the ability to do otherwise, they say, but moral responsibility does not. We can thus be morally responsible even if we do not have the ability to do otherwise – that is, even if determinism is true.

Compatibilist theories reject Claim (2) or Claim (4). **Libertarian theories**, on the other hand, reject Claim (3) (Section 13.7). There can be free will and moral responsibility, they say, even if determinism is false. **Simple indeterminist theories** claim that free actions are completely uncaused events. We can still be held morally accountable for our actions, however, because moral responsibility does not require that we actually bring about the actions we are responsible for. Unlike simple indeterminist theories, **causal indeterminist theories** deny that free actions are entirely uncaused. Free actions are caused by antecedent events, but they are caused in a way that is not deterministic. Antecedent events do not determine their results; they instead make their results more probable – they influence events without determining them. Finally, like causal indeterminist theories, **agent causal theories** claim that free actions have causes. They are nevertheless not caused by antecedent events in the way causal indeterminists claim; they are instead caused in a special kind of way by agents themselves. There is a special kind of causal relation in other words – *agent causation* – that differs from event causation, and that is responsible for bringing about actions. Finally, some libertarian theories deny the possibility of providing an informative positive account of how free will is possible. The eighteenth-century philosopher Immanuel Kant, for instance, argued for a position like this. Like other libertarians, Kant accepted claims (4) and (5), and thus concluded that free will must exist. He nevertheless

denied that we could ever understand how free will is possible. Our ways of conceiving the world, he said, are limited in such a way that we will never be able to understand exactly how free will is possible.

In addition to the foregoing theories, there are also some that deny the existence of free will and moral responsibility (Section 13.8). **Hard determinism** endorses Claim (2) together with determinism. Hard determinists, in other words, claim that free will and determinism are incompatible, and that determinism is true. From these two claims it follows that free will does not exist. **Hard incompatibilists**, on the other hand, believe that the truth or falsity of determinism is an empirical matter – something it is up to scientists not philosophers to determine. Consequently, hard incompatibilists do not take a stand on the truth or falsity of determinism. Like hard determinists, they endorse Claim (2), and like hard determinists they also reject the claim that we are free or morally responsible, but unlike hard determinists their reasons for rejecting the existence of free will or moral responsibility are not based on determinism, but on independent grounds.

In the sections that follow, we consider the foregoing theories in greater detail along with some of the arguments for and against each.

13.3 Classic Compatibilism

Classic compatibilists claim that the truth of determinism is compatible with the existence of free will; they thus try to solve the problem of free will and determinism by denying Claim (2). Recall the argument in favor of (2). Freedom requires the ability to do otherwise, it says, but an agent has the ability to do otherwise only if the agent has the ability to bring about more than one possible state of affairs, and determinism rules that out. Classic compatibilists agree that free will requires the ability to do otherwise, but they deny that the ability to do otherwise requires the ability to bring about more than one possible state of affairs. Instead, they say, free will requires only that people be able to do what they want in the absence of external constraints on their behavior – constraints such as being bound hand and foot, say, or being coerced at gunpoint, or suffering from some type of physical disability. So long as people act without being constrained in these kinds of ways, their actions are free, and to say they could do otherwise is to say that they would do otherwise if they wanted to.

Consider an example. According to classic compatibilists, to say that Madeleine broke the dishes freely is to say that Madeleine wanted to break the dishes, and was not forced to break the dishes by some type of external constraint: she was not coerced to break them at gunpoint, for instance, or forced to break them on account of some physical disability, say. In fact, Madeleine could have done otherwise, say classic compatibilists; that is, she would have done otherwise if she had wanted to. But since she wanted to break the dishes, she broke the dishes in

fact, and she did so freely. Classic compatibilists thus understand free will and the ability to do otherwise in terms of the ability to do what we want. They deny that the ability to do otherwise requires the ability to bring about different possible states of affairs.

If classic compatibilists are right, and free will requires only the ability to do what we want, then free will is compatible with determinism, for doing what we want is compatible with our wants being fully determined. Madeleine's desire to break the dishes at time t , the anger that caused that desire, the action that resulted from it, the absence of external constraints on Madeleine's behavior – all of these are factors that could have been produced deterministically by antecedent conditions in conjunction with the laws of nature. If determinism is true, therefore, this has no bearing on whether or not Madeleine's action was free, and hence it does not affect whether or not we hold her accountable for breaking the dishes.

Compatibilism was a popular view in the seventeenth and eighteenth centuries and was endorsed by prominent philosophers such as Thomas Hobbes, John Locke, and David Hume. Since the latter half of the twentieth century, moreover, compatibilism has been the default position for many philosophers in the free will debate. Compatibilism has nevertheless not been accepted by everyone. The philosopher Immanuel Kant called it a "wretched subterfuge," and the theory has met with some stiff challenges. One of them is the consequence argument.

13.4 The Consequence Argument

The consequence argument purports to show that compatibilism is false. We are free only if we are in control of our actions – only if what we do is up to us. If determinism is true, however, what we do is not up to us, say exponents of the argument; our actions are not in our control. Why is that the case? Here is one version of the argument: (1) If determinism is true, then our actions are necessary consequences of the laws of nature together with events in the past – including events that occurred before we were born. But (2) we are not in control of the laws of nature; what the laws of nature are is not something that is up to us. Moreover, (3) we are not in control of events in the past either; changing the past is also something that is not up to us. But (4) if we are not in control of the laws of nature or the past, if neither of them is up to us, then we are not in control of our actions either; what we do is also not up to us. Consequently, if determinism is true we are not in control of our actions; what we do is not up to us. But again, since we are free only if what we do is up to us, it follows that if determinism is true, we are not free. Compatibilism must be false. Let us consider the argument in detail.

Premise (1) follows directly from the definition of determinism: if determinism is true, then for any given state of the universe at a time, there is exactly one

possible state that comes about as a result of the laws of nature together with antecedent conditions. Premises (2) and (3), moreover, appear to be empirical matters of fact: it is not within our power to change either the past or the laws of nature. The most controversial premise is thus (4). The argument for (4) depends on two further premises which the philosopher Peter van Inwagen has dubbed 'Rule Alpha' and 'Rule Beta':

Rule Alpha: What is necessarily the case is not in our control.

Rule Beta: If X is not in our control, and Y is a necessary consequence of X, and it is not in our control that Y is a necessary consequence of X, then Y is not in our control either.

Consider an example. It is necessarily the case that two and two equal four. According to Rule Alpha, therefore, it is not up to us that two and two equal four – the sum of two and two equaling four is not something in our control. Suppose, moreover, that four and four equaling eight is a necessary consequence of two and two equaling four, and that this too is not in our control; there is nothing we can do to change it. In that case, according to Rule Beta four and four equaling eight is not in our control either; it is not something that is up to us.

Consider now the implications of Alpha and Beta for the argument. According to Alpha, we are powerless to change what is necessary, and according to Beta, we are powerless to change the necessary consequences of what is necessary. Suppose, then, that we are powerless to change the laws of nature and events in the past. If determinism is true, our actions are necessary consequences of the laws of nature and events in the past, and this too is something we are powerless to change. In that case, however, it follows from Alpha and Beta that we are powerless to change our actions, and in that case, it follows that we are not free if determinism is true.

The consequence argument poses a serious challenge to compatibilism. There are nevertheless several responses to it. One of them targets Rule Beta. According to critics, Beta is true only if one accepts a particular interpretation of what it means to be in control – of what it means for something to be up to us. But compatibilists reject that interpretation. On a compatibilist interpretation, critics argue, Rule Beta ends up being false. As a result, the consequence argument succeeds only if its defenders can show that the compatibilist interpretation of control is false; otherwise the consequence argument commits the fallacy of begging the question – it assumes rather than proves the falsity of compatibilism. Consider an example of the kind of interpretation of control that is supposed to falsify Rule Beta.

To falsify Beta, an interpretation of control must imply three things:

- (i) The laws of nature are not in our control.
- (ii) The past is not in our control.
- (iii) Our actions are in our control.

According to Rule Beta if (i) and (ii) are true, then (iii) must be false. According to Rule Beta, (i) and (ii) are supposed to imply that our actions are not in our control. Consequently, if compatibilists can provide an interpretation of control according to which (i) and (ii) are true, and our actions are in our control, then compatibilists will have an interpretation of control according to which Rule Beta is false. Suppose, then, that compatibilists say that we are in control of our actions provided we can do what we want free of external constraints. This appears to provide them with an account that satisfies statements (i)–(iii). Statement (i), for instance, says that the laws of nature are not in our control. According to the compatibilist interpretation of control, to say that the laws of nature are not in our control is to say that we could not change the laws of nature if we wanted to. Since there is in fact no circumstance in which we could change the laws of nature – even if we wanted to – Statement (i) ends up being true. The same is true of Statement (ii), the claim that events in the past are not in our control. On the compatibilist interpretation, to say that past events are not in our control is to say that we could not change those events even if we wanted to. Since we cannot in fact change past events, Statement (ii) ends up being true on the compatibilist interpretation as well. Consider finally Statement (iii), the claim that our actions are in our control. On the compatibilist interpretation, to say that our actions are in our control is to say that we could change our actions if we wanted to – that we would act differently if we wanted to. But surely, say compatibilists, this statement is true as well; surely if we had wanted to act differently on a given occasion, and there were no external constraints on our behavior, we would have acted differently. In that case, however, the compatibilist interpretation implies that statements (i)–(iii) are all true. But if (i)–(iii) are all true, then Rule Beta must be false, say critics, for Rule Beta says that if (i) and (ii) are true, (iii) must be false.

Because compatibilists can interpret control in this way, say critics, the consequence argument is not decisive. Defenders of the argument must show that the compatibilist interpretation is false; otherwise the argument simply begs the question against compatibilists. Defenders of the consequence argument have in fact advanced arguments to this effect. The compatibilist interpretation of control is false, they say. The reason is that there are cases in which the compatibilist interpretation implies that people are in control of their behavior when it is evident that they are not. Consider an example.

J. B. Watson, the founder of methodological behaviorism (Section 5.1), once conducted an experiment on an infant approximately 11 months old known as ‘Little Albert’. In the experiment (which would now be condemned as unethical) Little Albert was presented with a white laboratory rat. He initially showed no fear of the rat, but then Watson and his assistants produced a loud noise behind

the child's back by striking a steel bar. The child was startled by the noise, and became afraid. After the procedure was repeated several times, Little Albert began to show signs of fear at the mere appearance of the rat – and not just the rat but furry things in general: a rabbit, a teddy bear, a fur coat, even a Santa Claus mask made of cotton balls.

Imagine now that Little Albert grew to adulthood, and that in the course of his life whenever faced with choices between something furry and something not furry (a pet dog versus a pet lizard, a fuzzy wool hat versus a felt hat, and so on) Albert invariably chose the non-furry item. Were Albert's choices in his control in these cases? It seems evident that they were not, say critics of compatibilism. The scars left by Albert's childhood trauma made it impossible for him to desire furry things; he was forced to desire non-furry things, not by external constraints but by internal psychological ones. Choosing these things was thus not in his control; he did not have the ability to do otherwise. Yet the compatibilist interpretation gives the wrong result. On the compatibilist interpretation, say critics, choosing non-furry things was in Albert's control, for on the compatibilist's interpretation, to say that Albert's choices were in his control is to say simply that Albert would have chosen fur if he had wanted to and there were no external constraints on his behavior. But in the scenario we are imagining that is clearly the case: if indeed Albert had wanted furry things and there were no external constraints on his behavior, he would have chosen furry things. Therefore, on the compatibilist interpretation, Albert's behavior was in his control. But since Albert's choices were clearly not in his control, since he clearly lacked the ability to do otherwise, say critics, it follows that the compatibilist interpretation must be false.

Challenging Rule Beta is not the only way compatibilists have responded to the consequence argument. Another compatibilist response denies that free will and moral responsibility require the ability to do otherwise. This is the kind of response preferred by many contemporary compatibilists.

13.5 Contemporary Compatibilism I: Hierarchical and Capacity-Based Theories

Contemporary compatibilist theories differ from classic compatibilist theories in the way they approach the consequence argument. One way of defining free will in the context of the free will and determinism debate is to say that it is the kind of control over one's actions that is sufficient for genuine moral responsibility – the sort, in other words, that justifies us in evaluating people's behavior as good or bad, right or wrong. Exponents of the consequence argument assume that moral responsibility requires the ability to do otherwise – a claim the philosopher Harry Frankfurt has called the **Principle of Alternative Possibilities (PAP)**. As we saw in Section 13.3, classic compatibilists try to accommodate PAP. They argue that they can provide a compatibilist-friendly account of the ability to do otherwise,

an account compatible with determinism. Frankfurt and the contemporary compatibilists influenced by him have taken a different approach. Instead of offering compatibilist-friendly interpretations of PAP, they reject PAP altogether. PAP, they say, is false.

Critics of PAP argue against it by appeal to two kinds of examples. The first are called *character examples*. They involve character traits that apparently make it impossible for the people having them to do otherwise, and yet those traits and the actions issuing from them can still be evaluated as good or bad, right or wrong. Suppose, for instance, that Alexander cultivates a compassionate character that makes it impossible for him to harm innocent people. Harming an innocent person becomes something it is impossible for him to do. If Alexander is presented with an opportunity to harm an innocent person, even in exchange for an enormous financial reward, he steadfastly refuses, and his compassionate character prevents him from doing otherwise. Despite that, we can still evaluate his refusal to inflict harm on innocent people in moral terms. Hence, compatibilists conclude, moral responsibility does not require the ability to do otherwise.

Frankfurt himself, however, has advanced a different kind of example against PAP. Consider a case involving what is sometimes called a *Frankfurt controller*. Imagine that a mad neuroscientist implants a device in Gabriel's brain so that he is able to alter Gabriel's behavior with the push of a button. For various reasons, however, the neuroscientist does not want to interfere with Gabriel's behavior unnecessarily, and so he lets Gabriel do whatever he wants to do so long as it coincides with what he (the neuroscientist) wants Gabriel to do. If, for instance, the neuroscientist wants Gabriel to vote for political party A, and Gabriel plans on voting for A on his own, the mad neuroscientist will not intervene. It is only when it is clear that Gabriel is going to do something other than what the neuroscientist wants him to do, that the neuroscientist will intervene and alter Gabriel's behavior. In this case, for instance, the neuroscientist will intervene only if it becomes clear that Gabriel is about to vote for political party B.

Suppose now that Gabriel acts in a way that is morally praiseworthy or blameworthy. He does so, moreover, not because of the neuroscientist's intervention, but on his own initiative. The mad neuroscientist wants Gabriel to help an elderly woman carry her groceries, for instance, but Gabriel decides on his own to help the woman carry her groceries without the neuroscientist intervening. Or conversely, the mad neuroscientist wants Gabriel to steal something, but Gabriel decides on his own to steal it without the neuroscientist intervening. In these cases, it seems clear that we can hold Gabriel morally responsible for what he does, that his behavior is subject to moral praise and blame. Yet it is equally clear that in these cases Gabriel cannot do otherwise since the mad neuroscientist will not allow that to occur. As a result, say compatibilists, PAP must be false; moral responsibility does not require the ability to do otherwise.

Frankfurt-type examples purport to show that agents like Gabriel can be held morally accountable for their actions even though they lack the ability to do

otherwise. Many contemporary compatibilist theories are based on this idea. They nevertheless disagree in their positive accounts of moral responsibility. If free will and moral responsibility do not require the ability to do otherwise, what do they require? Frankfurt himself endorses a hierarchical theory of free will and moral responsibility.

Hierarchical theories understand freedom and responsibility in terms of a hierarchy of **higher-order** and **lower-order desires**. A higher-order desire is a desire for other desires. We can desire things: food, drink, sex, money, health, and so on. But we can also desire to have certain desires. Suppose, for instance, that I desire to be a generous person. Being a generous person involves having certain kinds of desires – generous people desire to help other people, for instance. Consequently, if I desire to be a generous person, I desire to have other desires; in particular, I desire to have a desire to help other people. A desire to help other people is a first-order desire, so my desire to be a generous person is a second-order desire; it is a desire to have certain first-order desires.

People's higher-order and lower-order desires can be in harmony. Truly generous people, for instance, not only want to help others; they also want to be the sort of people who want to help others. But higher-order and lower-order desires can also conflict. Consider an unwilling drug addict. The unwilling addict has a first-order desire to take the drug, but he does not want to be an addict; he does not want to be the sort of person whose desire for the drug influences or governs his behavior. His second-order desire not to be an addict is thus in conflict with his first-order desire to take the drug.

Hierarchical theories understand free will in terms of relations between higher-order and lower-order desires. The addict's behavior is compulsive; he cannot control his first-order desires. In particular, his *effective* first-order desires, the ones that influence or govern his behavior, are not the first-order desires he wants to be effective. The addict's behavior is therefore not free. If the addict were in control of his first-order desires, if he could make his effective first-order desires conform to his second-order desire not to be an addict, then according to hierarchical theories his behavior would be free. According to hierarchical theories, then, agents have free will exactly if they act in accordance with the first-order desires that they want to be effective – if, in other words, their effective first-order desires, the ones that actually influence or govern their behavior, conform to a second-order desire for these first-order desires to be the effective ones. If free will consists in this kind of harmony between higher-order and lower-order desires, say exponents of hierarchical views, then free will is compatible with determinism, for all of an agent's desires – first- and second-order – might be determined.

There are several objections to hierarchical theories of free will. First, according to hierarchical theories, free will consists in having our lower-order desires conform to our higher-order desires. Free will, however, is supposed to be the kind of control over behavior that grounds moral responsibility, and there are cases, say critics, in which we hold people morally accountable for their behavior even

though their lower-order desires do not conform to their higher-order desires. In addition, they say, there are cases in which we do not hold people morally accountable for their behavior even though their lower-order desires do conform to their higher-order desires.

An example of the first type of case involves weakness of will, a phenomenon many philosophers refer to by its Greek name *akrasia*. Imagine that I am a compulsive cookie-eater. Alexander bakes some cookies and I begin gluttonously devouring them. Just before I eat the last cookie, however, Alexander snatches it from my greedy hands and tells me that I cannot eat it because he is saving it for Xavier, the neighbor who generously donated the flour for his cookie-baking endeavor. Alexander leaves the room and I stare at the lone cookie on the plate. I want very much to eat it, but I do not want to be a gluttonous and ungracious person, and for that reason I do not want my desire for the cookie to govern what I do. But I eat the cookie anyway in spite of my second-order desire not to be governed by my first-order desire for it. In this case, my first-order desire fails to conform to my second-order desire. It seems evident, however, that I can be held accountable for my behavior – that my behavior in this case is morally blameworthy. And if I can be held morally accountable for my behavior, this suggests that my behavior is free since free will is supposed to be the kind of control over behavior that grounds moral responsibility. Hierarchical theories suggest the opposite however. They suggest that my behavior is not free since my first-order desire for the cookie does not conform to my second-order desire not to be an ungracious glutton, and, according to them, free will consists in lower-order desires conforming to higher-order ones. This type of case thus suggests that hierarchical theories give the wrong account of free will.

The same is true of the second type of case. These are cases in which an agent's first-order desire conforms to a second-order desire, but the second-order desire was formed under compulsion or some other freedom-undermining condition. Imagine, for instance, that I am kidnapped and my captors regularly inject me with a highly addictive drug so that I eventually become addicted. In addition, my captors manipulate my brain states so that I actually want to become an addict. In that case, my first-order desire to take the drug conforms to my second-order desire to be an addict, and in this case, say critics, it seems that I cannot be held accountable for any drug-seeking behavior that results. That behavior is not morally blameworthy because my second-order desires were subject to external manipulation that undermined my free will. Once again, however, hierarchical theories suggest exactly the opposite. Because my first-order desire for the drug conforms to my second-order desire to be an addict, they suggest that I still act freely since my first-order desire for the drug conforms to my second-order desire to be a drug addict, and they take free will to consist in the conformity of lower-order desires to higher-order ones. So like the previous example, say critics, this one suggests that hierarchical theories provide the wrong account of free will.

In addition, say critics, this example highlights another worry about hierarchical theories: the worry that they must postulate an infinite number of higher-order desires to give an adequate account of free will. The kidnapping example suggests that in order for my behavior to be free, not only must my first-order desires be formed apart from any freedom-undermining manipulation or compulsion, but my second-order desires must be formed apart from any freedom-undermining manipulation or compulsion as well. But the same point can be made with regard to third- or yet higher-order desires. If I am manipulated into desiring that the kidnappers should manipulate my brain states, or if I am manipulated into desiring that I be manipulated into desiring that kidnappers should manipulate my brain states, my freedom is undermined as well. In order to define free will hierarchically, therefore, it seems that hierarchical theories must stipulate that an agent's desires not be subject to coercion or manipulation not only at the first-order or the second-order, but also at the third-order, and the fourth-order, and the fifth-order, and so on *ad infinitum*. But postulating an infinite number of orders like this is absurd, say critics, and this too suggests that hierarchical theories give the wrong account of free will and moral responsibility.

A different compatibilist approach to free will denies that our desires and other motivational states must be arranged hierarchically to account for free will and moral responsibility. **Capacity-based theories** of free will and moral responsibility take motivational states to be tied to the exercise of various psychological capacities. People act freely, they say, when their behavior arises from exercising capacities of the right sorts – capacities for rational decision-making, say.

Exponents of a capacity-based theory of free will might claim, for instance, that free action involves acting in accordance with reasons – in accordance with what an agent believes to be best, and not merely with what the agent's appetites move him or her to do. On this kind of account, it is evident why my gluttonous cookie-eating is not really free. Free behavior is behavior that is governed by my beliefs about what is best. I believed it would be best not to eat the cookie, and acted contrary to that belief. Consequently, says the account, my behavior was not free but compulsory. It resulted from the exercise of my appetite for cookies alone, not from the exercise of my appetite moderated in light of my beliefs about what was best. Notice that if free will consists in the exercise of capacities in this way, then free will is compatible with determinism since the exercise of an agent's capacities might be determined.

Exponents of capacity-based theories of free will nevertheless disagree about which capacities are relevant to freedom. Like the capacity-based account just described, for instance, the philosopher Gary Watson distinguishes between our *values*, on the one hand, and our *desires* or *passions*, on the other. The former, he says, but not the latter are based on reasons about what it is best for us to do. We act freely if the behavior motivated by our desires or passions corresponds to the behavior dictated by our values. The philosopher Susan Wolf, on the other hand, claims that free will consists in the ability to act in accordance with *the True and*

the Good, and so people can be held morally accountable for their behavior only if they have this capacity.

One worry about capacity-based accounts like Watson's is that they appear to give us the wrong results when it comes to cases like weakness of will. According to those accounts, for instance, my cookie-eating was not free but compulsory. If my cookie-eating was not free, however, then I cannot be held accountable for it, and yet that seems wrong, say critics. Surely I can be held accountable for my behavior in that case. Capacity-based accounts like Susan Wolf's face an analogous worry. Imagine, for instance, a completely vicious person, a moral monster, who feels no remorse or regret even at committing the most heinous acts of theft, rape, murder, or violence – imagine, in other words, a person who has entirely lost the capacity to act in accordance with the True and the Good, as Wolf puts it. According to Wolf's theory, the moral monster is not free. If free will goes together with moral responsibility, however, if it is the kind of control over behavior that grounds moral evaluations, then it seems we cannot hold the moral monster accountable for his or her behavior. But again, say critics, that seems wrong. Surely we can hold moral monsters accountable for what they do. Hierarchical and capacity-based theories of free will remain controversial.

13.6 Contemporary Compatibilism II: Reactive Attitude Theories and Semicompatibilism

Reactive attitude theories ground moral responsibility in our social practices. The practice of moral evaluation involves attitudes, feelings, and intentions that people direct toward each other and themselves. People react to our attitudes toward them and our evaluations of them, and we in turn react to their attitudes toward us and their evaluations of us. Our reactive attitudes toward each other include gratitude, resentment, anger, forgiveness, love, and esteem. According to reactive attitude theories, we cannot help but have these attitudes toward each other; we cannot have ordinary interpersonal relationships without them. Sometimes we put these attitudes to one side and deal with people on a more or less impersonal level. Sometimes, for instance, we do not deal with difficult individuals on a personal level, as suitable targets for reactive attitudes, but on an impersonal level as problems to be solved or forces to be managed. When we put our reactive attitudes aside, and treat people impersonally in this way, it is not because we are convinced that determinism is true, but rather because we are convinced that there are circumstances that, for instance, diminish someone's capacity to be held accountable for his or her behavior – circumstances such as mental illness, fatigue, stress, or personal hardship. Conversely, when we treat people as ordinary targets of reactive attitudes, we do so not because we are convinced that determinism is false, but rather as a function of our more general participation in human social

practices – practices that we cannot help participating in because they satisfy deep humans needs for love, esteem, friendship, and the like. As a result, say exponents of reactive attitude theories, moral evaluation and hence moral responsibility do not depend in any way on the truth or falsity of determinism; they are instead grounded in human social practices that we must perforce engage in. There cannot be any serious question, therefore, of whether or not people are morally responsible for their actions, say reactive attitude theorists, and for the same reason there cannot be any serious question of whether or not moral responsibility is compatible with determinism. We can and do participate in moral evaluation independent of our beliefs about the truth or falsity of determinism. We can be morally responsible for our actions, then, even if determinism is true.

One objection to reactive attitude theories claims that the truth or falsity of determinism does in fact have a bearing on our practice of moral evaluation. Consider an example proposed by the philosopher Gideon Rosen. Imagine that Judas conspires to deliver Jesus to the Romans out of greed and envy, says Rosen. It seems that we are justified in blaming Judas for his behavior. But, says Rosen,

[S]uppose the whole thing was a setup. God's plan for salvation required that Jesus be betrayed, so [God] deliberately arranged the initial state of the universe and the laws of nature in such a way that Judas would betray him as he did with probability 1. That is to say, God saw to it that Judas would not exercise his capacity to do the right thing ... by seeing to it that it would be physically impossible for him to exercise this capacity in the circumstances. When we hear this it is hard not to be shaken by the initial conviction that Judas is not responsible for his act. It's not just that we come to think that God is *also* responsible. Something in the story tends to absolve Judas.¹

If God engineered the universe so that Judas could not have exercised his capacity to do the right thing, then there is a sense in which Judas's responsibility for his behavior is diminished. This poses a challenge to reactive attitude theories, for it suggests that determinism is relevant to moral responsibility, that our ordinary evaluative practices depend to some extent on the belief that people's behavior is not determined. Contrary to reactive attitude theories, therefore, the Judas example suggests that our ordinary evaluative practices depend on beliefs about the truth or falsity of determinism, that the truth or falsity of determinism is indeed relevant to the way we assign moral responsibility.

Like other compatibilists, *semicompatibilists* claim that moral responsibility is compatible with determinism. Unlike other compatibilists, however, semicompatibilists are willing to admit that free will is not compatible with determinism. If determinism is true, they say, there can be no free will. There nevertheless can be moral responsibility. According to semicompatibilists, in other words, moral responsibility does not require free will; in particular, it does not require the ability to do otherwise; it instead consists in having the right kind of control over our actions.

The most sophisticated semicompatibilist theory has been defended by John Martin Fischer and Mark Ravizza. Fischer and Ravizza call the control required for moral responsibility *guidance control*. Recall the Frankfurt controller described in Section 13.5. Gabriel is faced with a choice of voting for political party A or political party B. The mad neuroscientist wants Gabriel to vote for A, and will prevent Gabriel from voting for B if it becomes evident that this is what Gabriel is going to do. It is thus impossible for Gabriel to do otherwise. Suppose, however, that Gabriel decides to vote for A on his own without the neuroscientist intervening. It seems clear, say semicompatibilists, that we can hold Gabriel responsible for his action in this case even though he lacks the ability to do otherwise. Why is that the case? The reason, they say, is that Gabriel's action was still in his control; in particular, he exercised guidance control over it, and this is the kind of control that enables us to hold agents responsible for what they do even if they lack the ability to do otherwise.

What exactly is guidance control? According to Fischer and Ravizza it has two components. First, the mechanism that produces the action must be *responsive to reasons*. Second, the agent must *take responsibility* for that mechanism; in particular, the agent must see him or herself as a target of reactive attitudes that are due to the actions produced by that mechanism. Let us consider these conditions one at a time.

For a mechanism to be responsive to reasons, it must not be a mechanism that produces actions compulsively, the way that, say, a drug addict might take a drug compulsively. If, for instance, Gabriel compulsively votes for political candidates that are listed on the left-hand side of the ballot, and political party A happens to be listed on the left-hand side of the ballot, then Gabriel's voting for party A is not responsive to reasons; it is not the sort of behavior that would be altered by, say, learning something about which policies political party A will support. Suppose, however, that Gabriel votes for party A because he wants lower property taxes, and believes party A will do more to lower property taxes than party B. In that case, the mechanism that produces his voting is responsive to reasons. If, for instance, Gabriel came to believe that party B was more likely to lower property taxes than party A, then Gabriel would vote for party B instead – a change in Gabriel's reasons would result in a different action.

In addition to reasons-responsiveness, say Fischer and Ravizza, guidance control also requires that an agent take responsibility for the actions produced by a mechanism. For an agent to take responsibility in this sense is for the agent to recognize that he or she will be a target of reactive attitudes on account of the actions produced by that mechanism. If Gabriel, for instance, understands that people will hold him accountable for voting for A, that some will be pleased and others disappointed or angered by his vote, then Gabriel can be said to take responsibility for the mechanism producing his action.

If semicompatibilists are right about the requirements for moral responsibility, then it seems that moral responsibility is compatible with determinism, for agents

can take responsibility for actions that are produced by reasons-responsive mechanisms even if those mechanisms operate in a completely deterministic way. Critics of semicompatibilism might nevertheless wonder whether the semicompatibilist account captures everything involved in moral responsibility. Consider again Gideon Rosen's Judas example. It seems possible that Judas might exercise guidance control over his action – that his betrayal of Jesus might be produced by a reasons-responsive mechanism for which Judas takes responsibility. In that case, however, semicompatibilist theories face a problem analogous to the problem facing reactive attitude theories. Even though Judas exercises guidance control over his action, critics might say, there is still a strong intuition that his responsibility for that action is diminished. There is an intuition, in other words, that the truth of determinism plays an important role in the way we assign moral responsibility – a role not included in the idea of guidance control.

Each compatibilist theory thus has its detractors. In addition to criticisms that target particular compatibilist theories, moreover, opponents of compatibilism have also criticized the contemporary compatibilist paradigm in general. Recall that many contemporary compatibilist theories are based on a rejection of PAP, and that their exponents reject PAP on the basis of Frankfurt-type examples. Critics of compatibilism nevertheless argue that those examples are not decisive. Consider one argument to this effect; we can call it the *indeterminist objection* to Frankfurt-type examples.

According to the indeterminist objection, Frankfurt controllers can exist only in a deterministic universe. Consider again the mad neuroscientist who looks to control Gabriel's behavior (Section 13.5). He interferes with Gabriel's behavior only when it becomes evident to him that Gabriel is going to act contrary to the way he wants Gabriel to act. But how is the neuroscientist supposed to know whether Gabriel will act contrary to his wishes? Unless determinism is true, it is difficult to see how he could. To see this, imagine that determinism is false. In that case, even if the neuroscientist knows all of the events leading up to Gabriel's decision to vote for party A or party B, he still does not know whether Gabriel will vote for A or B, for if determinism is false, it is possible for the universe to go either way completely independent of any antecedent conditions. Two parallel universes that have exactly the same laws of nature and past events could still have different futures if determinism is false. Unless the universe is deterministic, therefore, no knowledge of prior conditions and the laws of nature will tell the mad neuroscientist how Gabriel will vote. In order for a Frankfurt controller to exist, therefore, determinism must be true.

This means, however, that when compatibilists endorse Frankfurt-type examples, they are implicitly endorsing determinism. But if that is the case, then Frankfurt-type examples do not succeed in refuting incompatibilist theories that claim determinism is false. They do not succeed, for instance, in refuting libertarian theories, compatibilists' main competitors. For compatibilists to insist that Frankfurt-type examples give us reason to think that libertarian theories are false

is to commit the fallacy of begging the question, for exponents of Frankfurt-type examples must assume that determinism is true, and that means they must assume that libertarianism is false. Exponents of Frankfurt-type examples do not prove that libertarianism is false, therefore; they instead take the falsity of libertarianism as an unstated and undefended assumption.

What does the indeterminist objection mean for the free will debate? According to critics of compatibilism, it means that Frankfurt-type examples do not provide decisive evidence in favor of compatibilism; in particular, they do not show that libertarian theories are false. Libertarians and others who reject determinism are thus free to reject Frankfurt-type examples, and with them one of the central pillars of contemporary compatibilist thinking.

13.7 Libertarianism: Simple Indeterminism, Causal Indeterminism, and Agent Causation

Libertarian theories claim that we can be free and responsible even if determinism is false. The burden for any libertarian theory is to explain how people are able to exercise control over their actions without antecedent conditions that determine their effects.

Recall that if determinism is false, then for a given state of the universe at a time, there is more than one possible state that can result. Recall, moreover, that if determinism is false, nothing can be responsible for bringing about one possible state as opposed to another. Suppose, for instance, that it is possible for Madeleine to break the dishes at time t or not to break the dishes at t , and that she breaks the dishes in fact. What explains this result? Why did one state of affairs result as opposed to the other? If determinism is true, there is a straightforward answer: Madeleine broke the dishes because that is what was determined by certain antecedent conditions in conjunction with the laws of nature. If determinism is false, however, the answer is no longer clear, for if determinism is false, there can be no factor that was responsible for determining which state among the range of possible alternatives actually occurred. There is no causal factor, no law or antecedent condition that brought about Madeleine's action. Neither Madeleine nor her thoughts, choices, reasons, or motives were responsible for her act of breaking the dishes. Recall, finally, the problem this poses for moral responsibility. If nothing was responsible for Madeleine's act of breaking the dishes; if that action, and actions in general merely happen at random, then it looks like we cannot hold people accountable for what they do since their actions are not ultimately up to them. Ultimately whether or not Madeleine broke the dishes was not in her control; ultimately, whether the universe included the breaking of the dishes or not was something that happened at random. None of us, then, is ultimately responsible for what we do; ultimately, none of our actions is in our control. But

if none of our actions is in our control, then we cannot be held morally accountable for what we do. Consequently, if determinism is false, there can be no moral responsibility. This is one of the problems libertarian theories face. Call it *the control problem*. Consider another.

Imagine that Madeline is faced with a choice: she can accept a job offer from law firm A or a job offer from law firm B. Suppose moreover that firm A has offered Madeleine more money than firm B and more flexible hours. It is also an easier commute, and it sponsors pro bono work that will enable Madeline to feel good about what she does. These seem like very good reasons for Madeleine to prefer law firm A to law firm B, and we might imagine that there are no other reasons that would favor law firm B. The issue thus seems settled: there are decisive reasons to accept the offer from firm A instead of firm B. Notice, however, that if determinism is false the issue is not settled, for if Madeleine's choice is not determined by anything, then it is not determined by her reasons. Madeleine could thus choose firm B irrespective of the very good reasons she has for choosing firm A. This scenario highlights another problem facing libertarian theories. If Madeleine chooses firm B, her choice is completely inexplicable. Given the reasons favoring firm A, there is no way we can explain Madeleine's decision. If determinism is false, it is not possible to explain human decisions and actions; in particular, it is not possible to explain people's behavior by appeal to reasons. If actions are not determined by antecedent conditions, then they are not determined by people's reasons. If determinism is false, therefore, reasons no longer have any explanatory power. Call this *the reason problem*.

Consider a final problem facing libertarians. If what happens in the universe is not determined, if it happens instead by chance, then it looks like moral evaluations of behavior are not justified. Suppose, for instance, that choosing law firm A would be morally praiseworthy, and choosing law firm B would be morally blameworthy. Since Madeline's choice of one firm over the other is ultimately not something in her control, since it is instead something that happens by chance, we cannot praise her or blame her for choosing one firm as opposed to the other, for we cannot praise or blame people for things that are beyond their control. If determinism is false, however, then Madeleine's decision is ultimately beyond her control. We are not justified in either praising or blaming her for her decision any more than we are justified in praising her or blaming her for the weather, for if determinism is false, neither her decision nor the weather is within her control. If determinism is false, therefore, we can never be justified in holding people morally accountable for what they do. Call this *the moral justification problem*. It is also sometimes called *the problem of luck* since it claims that it is ultimately a matter of luck whether an agent's behavior counts as praiseworthy or blameworthy if determinism is false.

These are some of the problems facing libertarian theories. Different libertarian theories try to address them in different ways. Those theories are of at least three different kinds. *Simple indeterminist theories* claim that actions are completely

uncaused. *Causal indeterminist theories*, on the other hand, claim that actions are caused by antecedent events, but that they are caused in ways that are merely probabilistic not deterministic. Finally, *agent-causal theories* claim that actions are produced by a special kind of causal relation unique to agents – agent causation. Let us consider these theories in order.

Simple indeterminist theories claim that actions are completely uncaused, but that this is compatible with the explanatory power of reasons and the justifiability of moral evaluations. The philosopher Carl Ginet defends a theory along these lines. According to Ginet, some of the events in our lives have an “actish phenomenal quality”: when we experience them they seem to us as if we are acting and not merely being acted upon. On Ginet’s view, these events are actions. Among actions, the free ones are the ones that we do for reasons and that are also undetermined. For an action to be done for a reason, according to Ginet, is for that action to be explainable by appeal to the agent’s beliefs, desires, intentions, or other mental states. Free actions are thus explainable by appeal to reasons, and according to Ginet that means our moral evaluations of them can be justified since moral evaluations usually depend on assessments of people’s reasons for doing what they do. There are nevertheless several objections to simple indeterminism.

One objection concerns how reasons are supposed to explain actions on the indeterminist view. The philosopher Donald Davidson, for instance, once argued that reasons had to be causes, for otherwise there would be no way of distinguishing the reason the agent performed an action from just any old reason for performing an action of that sort (Section 7.6). According to Davidson, then, reasons must be causes. Simple indeterminist views deny that this is the case however. How then do reasons explain actions if not by causing them? According to Ginet, reasons explain actions by entering into our intentions to perform them. Madeleine’s reasons for preferring law firm A to law firm B, for instance, explain her action because in choosing the one as opposed to the other, Madeleine intends to satisfy a desire for better pay, more flexible hours, a shorter commute, and conscience-satisfying work. Critics of Ginet’s view nevertheless argue that we often have reasons for performing actions that never enter into our intentions. Sometimes, for instance, we have subconscious motivations, or motivations that we repress psychologically, and hence that never enter into our intentions to perform various actions. Consequently, say critics, Ginet does not offer a satisfactory account of rational explanation.

Moreover, say critics, even if simple indeterminists were able to solve the reason and moral justification problems, they would still have a problem with control. Because agents do not cause actions on simple indeterminist views, they are ultimately not in control of the actions they perform. But if agents are not in control of the actions they perform, then it is difficult to see how they could really be held accountable for those actions whatever their reasons for performing them.

A similar problem confronts causal indeterminist theories. Unlike simple indeterminist theories, causal indeterminist theories claim that actions have causes. Those causes, however, do not determine their effects the way some causal relations do, say causal determinists; they merely make those effects more probable. According to causal indeterminists, for instance, Madeleine's desire to break the dishes does not produce her act of breaking the dishes with a probability of 1; it merely makes that act more likely to occur, and there remains a chance she won't break the dishes even if she wants to. Madeleine's desire thus influences her behavior without determining it; her desire causes her action indeterministically.

Causal indeterminist theories do more to address the control problem than simple indeterminist theories. Critics argue, however, that indeterministic causation still does not provide the kind of control that is needed for moral responsibility. Even though Madeleine's desire makes her action more probable, there is still a chance the action will not occur. Suppose, for instance, that Madeleine's desire makes it probable that she will break the dishes with a probability of .99. Given that there is still a chance Madeleine will not break the dishes, what ultimately tips the balance in favor of Madeleine breaking them as opposed to not? If there is no balance-tipping factor, if it is merely chance, then once again there is a problem with control, and hence a problem with moral responsibility.

One type of libertarian view tries to address this problem by combining indeterministic causation with deterministic causation. Suppose, for instance, that Madeleine's desires and intentions are formed at least in part by chance. In the process of deciding what to do, a range of diverse factors enter her mind – many of them completely at random. This is the indeterministic component of the view. At a certain point in the decision-making processes, however, the factors that are going to produce Madeleine's actions become fixed; they become *the* reasons Madeleine acts as she does, and they produce Madeleine's action deterministically. This is the deterministic component of the view. A libertarian view that combines deterministic and indeterministic elements in this way goes some of the way to addressing the control problem. Because the agent's reasons deterministically produce the action, the agent can be said to control the action. Yet because the agent's reasons are formed partly by chance, the view is not altogether deterministic, and this will perhaps be enough to give libertarians what they want.

Many libertarians are nevertheless dissatisfied with this approach to the control problem, and try to supply the control needed for moral responsibility by appeal to something else. The philosopher Robert Kane, for instance, claims that the control needed for moral responsibility is supplied by an *effort of will* on the part of the agent. An effort of will is sufficient to supply the kind of control needed for moral responsibility, Kane argues, because when agents make efforts of will, we typically hold them accountable for what they do even if the results of those efforts are undetermined. Kane illustrates this idea with an analogy. Suppose that Xavier tries to shatter a glass tabletop by striking it with his fist. Suppose, moreover, that it is not determined whether or not he will succeed. If Xavier does

succeed in fact, it is clear that we will hold him morally accountable for breaking the tabletop. It seems, therefore, that the indeterminacy of an effect does not stop us from assigning moral responsibility for that effect when the agent is trying to produce it – when, in other words, the agent makes an effort of will.

Another similar approach to the control problem postulates a special kind of causal relation – *agent causation*. Causal indeterminist theories of the sort we have described so far are sometimes called forms of *event-causal indeterminism* to distinguish them from forms of *agent-causal indeterminism*, libertarian views that endorse agent causation. According to agent-causal theories, actions are subject to a special kind of causal relation; they are caused directly by agents not by events, and for that reason agents exercise direct control over them, and are thus directly responsible for their occurrence. According to agent-causal theories, in other words, actions are distinguished from other events because they are caused directly by agents. The agent-causal relation thus supplies agents with the kind of control that is needed for moral responsibility, say agent-causal theorists, and yet it is not a deterministic causal relation.

There are at least two objections to agent-causal views. First, critics argue that agent causation is not a special type of causal relation, but just a species of event causation. After all, critics argue, it is not just any feature of the agent that is responsible for producing the agent's actions, but the agent's reasons specifically. It is, for instance, Madeleine's desire for better pay, better hours, and a more satisfying job that causes her to choose law firm A as opposed to law firm B. It is, in other words, her having a desire at time *t* that causes her action. But Madeleine's having a desire at *t* is an event. It is an event, therefore, that causes Madeleine's action. According to critics, then, what agent-causal theorists take to be a special kind of causal relation appears to be merely a species of a more familiar type of causal relation, the kind that obtains between events. In response to this objection, some agent-causal theorists have been willing to grant that agent causation can be understood as a species of event causation provided that event causation is not understood in a way that is committed to determinism or that ends up undermining the agent's control over the action.

The second objection to agent-causal theories is more serious. It argues that the agent-causal solution to the problem of free will and determinism is devoid of any real content. According to agent-causal theorists, say critics, agent causation is supposed to be a special causal relation that satisfies two conditions:

- 1 It supplies agents with the kind of control needed for moral responsibility.
- 2 It is not a form of deterministic causation.

If there is a form of causation that satisfies conditions (1) and (2), then clearly it provides a compelling solution to the problem of free will and determinism. But why should we suppose there is such a form of causation? According to critics of agent causation, there is no good reason. Consider an analogy with the objection

to neutral monism discussed in Section 9.5. Neutral monists claim that everything is in itself neither mentalistic nor physicalistic but neutral. If neutral monists are right, their theory provides a unique and attractive solution to mind-body problems. The difficulty is that neutral monists have never been able to give an informative account of what neutral entities are. They claim that neutral entities are ones that can be described in both mental and physical terms – they give a definition, in other words, that describes the role neutral entities are supposed to play in their theory; what they do not give is any information that would enable us to determine which entities in the world, if any, actually occupy that role. Without an informative account of what neutral entities are, however, neutral monism remains merely an abstract possibility – a suggestion, not a genuine theory that can be evaluated alongside its competitors.

The argument against agent causation is analogous to the argument against neutral monism. Critics claim that agent-causal theorists have failed to give an informative account of what agent causation is; they have said that agent causation is a causal relation that satisfies conditions (1) and (2) – they have given a definition, in other words, that describes the role agent causation is supposed to play in agent-causal theory; what they have not given is any information that would establish that such a relation really exists. They have not given any reason to think that there really is a causal relation that satisfies conditions (1) and (2). Because of that, say critics, the agent-causal solution to the problem of free will and determinism remains a mere abstract possibility, not a genuine solution that can be evaluated alongside its competitors.

One way agent-causal theorists can respond to this argument is by appealing to an account of causation like the causal pluralism endorsed by hylomorphists (Section 10.7). Recall that hylomorphists base their account of causation on an account of explanation. Causal relations, they say, are explanatory relations, and causes are explanatory factors. When it comes to determining what causes and causal relations really exist, hylomorphists defer to the relevant empirical disciplines – to our best empirical accounts of how and why things behave as they do. Suppose, then, that our best explanations of human behavior include causal relations that satisfy conditions (1) and (2). In that case, the notion of agent causation would have real empirical content. Agent causation would not be a mere hypothetical postulate cooked up by libertarians to deal with the problem of control; it would instead be a kind of causal relation that factors into our best empirical explanations of human behavior.

One final libertarian theory worth mentioning is the kind of theory endorsed by the eighteenth-century philosopher Immanuel Kant. Kant claimed that moral responsibility requires free will in the libertarian sense, and concluded that because we are in fact morally responsible for our behavior, free will in the libertarian sense must exist. But Kant also denied the possibility of solving the problem of free will and determinism – of reconciling the existence of libertarian free will with either the truth or the falsity of determinism. Kant argued that there are absolute limits

on our ability to understand the world, and that when we try to exceed those limits, we encounter philosophical problems. The problem of free will and determinism is one of them. A real solution to the problem lies beyond our grasp, and it will forever lie beyond our grasp since it requires knowing how things are in themselves. On Kant's view, however, we can never know how things are in themselves because our understanding of things is entirely dependent on the structure of our minds. Because our minds are structured the way they are, we cannot help but understand the world in certain ways – in terms of time, space, and causation, for instance. Because our understanding of the world is limited to categories like this, we will never be able to grasp how beings like us might be capable of libertarian freedom. We can nevertheless be confident that libertarian freedom exists because we are morally responsible for our actions, and libertarian freedom is necessary for moral responsibility.

Like the other libertarian theories we have considered, Kant asserts the existence of libertarian freedom on the basis of our experience of moral obligation and moral evaluation. Unlike those theories, however, he denies the possibility of giving a positive account of libertarian freedom that reconciles it with the truth or falsity of determinism.

13.8 Hard Determinism and Hard Incompatibilism: Do Freedom and Responsibility Really Exist?

Consider finally views that reject the existence of genuine moral responsibility. The classic version of this type of view is *hard determinism*. Hard determinists endorse the conjunction of incompatibilism and determinism. They claim, in other words, that free will and determinism are incompatible, and that determinism is true. From these two claims it follows that free will does not exist. But if free will does not exist, and moral responsibility requires free will, it follows that moral responsibility does not exist either.

Many contemporary deniers of free will and moral responsibility are not hard determinists, however. They believe that the truth or falsity of determinism is an empirical matter – something it is up to scientists to determine not philosophers. Consequently, they do not take a stand on the truth or falsity of determinism as hard determinists do. Like hard determinists, they endorse incompatibilism, and also reject the claim that we are free, but their rejection of this claim is based not on determinism, but on independent grounds. This view is sometimes called **hard incompatibilism**. The label 'incompatibilism' indicates the view's rejection of compatibilist accounts of free will, and the adjective 'hard' indicates its rejection of libertarian accounts of free will in addition.

Philosophers who deny moral responsibility disagree about what implications their view has for our ordinary institutions and practices. Some claim the denial

of real moral responsibility would have no serious implications for our institutions. Call this a *deflationary* approach to denying moral responsibility since it tries to deflate the denial's practical implications. According to the deflationary approach, denying the existence of moral responsibility might require us to alter the ways we understand our institutions, or conceive of their purposes or functions within society, but we would not have to alter the institutions themselves in any significant way. Denying the existence of moral responsibility, for instance, would not require us to stop the practice of punishing people for wrongdoing. If there is no free will, say exponents of the deflationary view, then we might have to give up the idea that punishment is deserved in some deep moral sense, or that one of the functions of punishment is to provide retribution for harm done, but we do not have to get rid of the institution of punishment since it might still perform other functions in society. It might perform an important deterrent function, for instance, or help to reform or rehabilitate offenders.

Other philosophers, however, think that denying moral responsibility would have implications for our institutions and practices that would go beyond changing the way we conceive of them. We can call this a *revisionary* approach to denying moral responsibility. Consider love. Some revisionists claim that many people will need to revise their expectations about having loving relationships if there is no free will or moral responsibility. People who think that genuine loving relationships must be freely chosen in a libertarian sense, for instance, will have to revise their ideas about whether such relationships are really possible, and in that case, they will also have to revise any behavior that is directed toward achieving such relationships since it would be irrational to pursue a goal that was unattainable. Revisionists nevertheless disagree on the extent to which our institutions and practices will have to be revised if freedom and moral responsibility do not exist, and some of them find the prospects of these revisions so alarming that they advocate living with and promoting the illusion of freedom and moral responsibility.

What reasons are there, however, to think that morally responsibility does not really exist, that free will is at best an illusion? Here is one version of the kind of argument opponents of free will and moral responsibility sometimes offer:

- 1 You are responsible for your behavior only if you are responsible for all the factors that contributed to forming your character, the source of the reasons and motivations that govern your behavior.
- 2 It is not the case that you are responsible for all the factors that contributed to forming your character.

Therefore, you are not responsible for your behavior.

We have already considered some examples that support Premise (1). Consider the examples used to criticize hierarchical theories of free will. They suggest that

in order for an action to be free it must be the case not only that the action is in the agent's control but also that the desires or other mental states that motivated it be in the agent's control. Consider the kidnapping example discussed in Section 13.5. If my captors manipulate my brain and produce in me a desire to be a drug addict, then I cannot be held accountable for having that desire, and if I cannot be held accountable for having that desire, then I cannot be held accountable for the drug-seeking behavior in which I engage as a result. I can be responsible for my drug-seeking behavior, then, only if I am also responsible for the higher-order desires that motivate it. Examples like this suggest that to be responsible for our behavior we have to be responsible for the factors that shaped our desires and motivations.

Premise (2), on the other hand, is supported on empirical grounds. There appear to be many factors that contributed to forming our characters over which we exercised little or no control. You had little or no control over your parents' actions and attitudes, for instance, or the environment in which you were raised, or the people by whom you were surrounded, and so on. Because you were not responsible for these factors, say exponents of the argument, you were not responsible for forming your character.

Your character, however, is what governs your behavior; it is the source of your reasons and desires for acting in the ways you do. In order to be fully responsible for what you do, therefore, you would have to be fully responsible for forming your character. But you are not fully responsible for forming your character since many of the factors that went into forming it were completely beyond your control. Because we cannot blame you for the character you have, we cannot blame you for the actions you perform, and conversely, because we cannot praise you for the character you have, we cannot praise you for the actions you perform. You are, therefore, not really responsible for your actions.

Critics of the argument will have a difficult time attacking Premise (2). It is difficult to defend the claim that we are responsible for all the factors that contributed to forming our characters. Critics of the argument might nevertheless take aim at Premise (1). While it is true that moral responsibility requires that we have control over some of the factors that formed our characters, critics might say, it does not require that we have control over all of them. I might not be responsible for the way in which my parents acted, or the attitudes they had, or the environment in which I was raised, but I was and am responsible for other factors that contributed to my character. At a certain point in my life, I began making decisions on my own for reasons independent of the reasons my parents or other caregivers had. Because I had some control over these factors, say critics of the argument, I can be held accountable to some extent for the character traits that resulted. But if I can be held accountable to some extent for my character, I can be held accountable to some extent for the actions that issue from it. Premise (1) is therefore false. Arguments for and against the existence of free will and moral responsibility remain controversial.

Further Reading

There are several good anthologies on the problem of free will and determinism. They include Watson (2003), Kane (2002), and Ekstrom (2001). More advanced students might find the essays in Kane (2001) helpful as well. Kane (2005) provides a helpful general introduction to the free will debate.

Classic compatibilism was defended by Thomas Hobbes (1996 [1651]), John Locke (1959 [1690]), and David Hume (2007 [1748]). For more recent defenses of classic compatibilism see Smart (1961) and Nielsen (2002). Daniel Dennett (1978) provides a lively defense of compatibilist thinking.

Prominent defenders of the consequence argument include Peter van Inwagen (1983) and Carl Ginet (1990). Frankfurt's original argument against PAP is presented in Frankfurt (1969). The indeterminist objection to Frankfurt-type examples was originally advanced by Robert Kane (1985: 51). It has since been advanced by others including David Widerker (1995), Carl Ginet (1996), and Keith Wyma (1997). Peter van Inwagen (1983) criticizes Frankfurt-type examples in a different way. For more on Frankfurt-type examples, criticisms of them, and compatibilist responses to those criticisms see the essays in Widerker and McKenna (2003).

Harry Frankfurt develops his hierarchical account of free will in the first five essays in Frankfurt (1988). Gerald Dworkin (1988) defends a different hierarchical theory. Defenders of capacity-based compatibilist theories include Gary Watson (1975) and Susan Wolf (1980). P. F. Strawson (1974) defends a reactive attitude theory, as does R. Jay Wallace (1994). See Rosen (2002) for the Judas example, and Fischer and Ravizza (1998) for their defense of semicompatibilism.

Carl Ginet (1990) defends his simple indeterminist view of free will. Hugh McCann (1998) defends a similar view. Daniel Dennett (1984) articulates but does not endorse the kind of causal indeterminist view that combines deterministic with indeterministic elements. Alfred Mele (1995) does the same. Robert Kane (1985) defends a causal indeterminist view. He discusses the shattered tabletop example in Kane (1999).

The eighteenth-century Scottish philosopher Thomas Reid is considered the father of modern agent-causal theories. See, for instance, "Of the Liberty of Moral Agents," Essay IV of his "Essays on the Active Powers of the Human Mind," Part C of *The Works of Thomas Reid: Volume 1*. Contemporary defenders of agent-causal theories include Roderick Chisholm (2002), Richard Taylor (1992), Timothy O'Connor (2000), and Randolph Clarke (1993; 1996). Immanuel Kant (1993 [1785]) develops Kant's libertarian views on free will.

Hard determinism was defended in the eighteenth-century by the Baron d'Holbach (1999 [1770]), and it was defended in the twentieth-century by the philosopher Paul Edwards (2002). Galen Strawson (1986; 1994) advances an argument against moral responsibility. Other opponents of free will and moral responsibility include Ted Honderich (2002) and Derk Pereboom (2001). Both take a

largely deflationary view of rejecting free will and moral responsibility. Saul Smilansky (2000) is not so sanguine. He argues that we need to maintain the illusion of free will if we are to preserve our institutions and social practices.

Note

- 1 Gideon Rosen, 2002, "The Case for Incompatibilism," *Philosophy and Phenomenological Research* 64/3: 699–706, 703.

References

- CHISHOLM, RODERICK M. 2002. "Human Freedom and the Self." In *Free Will*, edited by Robert Kane, 47–58. Malden: Blackwell Publishers.
- CLARKE, RANDOLPH. 1993. "Toward a Credible Agent-Causal Account of Free Will." *Nous* 27/2: 191–203.
- CLARKE, RANDOLPH. 1996. "Agent Causation and Event Causation in the Production of Free Action." *Philosophical Topics* 24/2: 19–48.
- DENNETT, DANIEL CLEMENT. 1978. "On Giving Libertarians What They Say They Want." In *Brainstorms*, 286–99. Montgomery: Bradford Books.
- DENNETT, DANIEL CLEMENT. 1984. *Elbow Room: The Varieties of Free Will Worth Wanting*. Cambridge: MIT Press.
- D'HOLBACH, BARON. 1999 [1770]. *System of Nature*. Translated by H. D. Robinson and Alastair Jackson. Manchester: Clinamen Press.
- DWORKIN, GERALD. 1988. *The Theory and Practice of Autonomy*. New York: Cambridge University Press.
- EDWARDS, PAUL. 2002. "Hard and Soft Determinism." In *Free Will*, edited by Robert Kane, 59–69. Malden: Blackwell Publishers.
- EKSTROM, LAURA WADDELL. 2001. *Agency and Responsibility: Essays on the Metaphysics of Freedom*. Boulder: Westview Press.
- FISCHER, JOHN MARTIN, and MARK RAVIZZA. 1998. *Responsibility and Control: A Theory of Moral Responsibility*. New York: Cambridge University Press.
- FRANKFURT, HARRY. 1969. "Alternative Possibilities and Moral Responsibility." *Journal of Philosophy* 66: 829–39.
- FRANKFURT, HARRY. 1988. *The Importance of What We Care About*. New York: Cambridge University Press.
- GINET, CARL. 1990. *On Action*. New York: Cambridge University Press.
- GINET, CARL. 1996. "In Defense of the Principle of Alternative Possibilities: Why I Don't Find Frankfurt's Argument Convincing." *Philosophical Perspectives* 10: 403–17.
- HOBBS, THOMAS. 1996 [1651]. *Leviathan*. Edited by Richard Tuck. New York: Cambridge University Press.
- HONDERICH, TED. 2002. *How Free Are You?* 2nd edn. New York: Oxford University Press.
- HUME, DAVID. 2007 [1748]. *An Enquiry Concerning Human Understanding and Other Writings*. Edited by Stephen Buckle. New York: Cambridge University Press.

- KANE, ROBERT. 1985. *Free Will and Values*. Albany: State University of New York Press.
- KANE, ROBERT. 1996. *The Significance of Free Will*. New York: Oxford University Press.
- KANE, ROBERT. 1999. "Responsibility, Luck, and Chance: Reflections on Free Will and Indeterminism." *Journal of Philosophy* 96/5: 217–40.
- KANE, ROBERT. 2001. *The Oxford Handbook of Free Will*. New York: Oxford University Press.
- KANE, ROBERT. 2002. *Free Will*. Malden: Blackwell Publishers.
- KANE, ROBERT. 2005. *A Contemporary Introduction to Free Will*. New York: Oxford University Press.
- KANT, IMMANUEL. 1993 [1785]. *Grounding for the Metaphysics of Morals*. Translated by James W. Ellington. 3rd edn. Indianapolis: Hackett Publishing Co.
- LOCKE, JOHN. 1959 [1690]. *An Essay Concerning Human Understanding*. 2 vols. New York: Dover Publications, Inc.
- MCCANN, HUGH. 1998. *The Works of Agency: On Human Action, Will, and Freedom*. Ithaca: Cornell University Press.
- MELE, ALFRED R. 1995. *Autonomous Agents: From Self-Control to Autonomy*. New York: Oxford University Press.
- NIELSEN, KAI. 2002. "The Compatibility of Freedom and Determinism." In *Free Will*, edited by Robert Kane, 39–46. Malden: Blackwell Publishers.
- O'CONNOR, TIMOTHY. 2000. *Persons and Causes: The Metaphysics of Free Will*. New York: Oxford University Press.
- PEREBOOM, DERK. 2001. *Living without Free Will*. New York: Cambridge University Press.
- REID, THOMAS, WILLIAM HAMILTON, and DUGALD STEWART. 1872. *The Works of Thomas Reid*, Vol. 1. 6th edn. Edinburgh: Maclachlan and Stewart.
- ROSEN, GIDEON. 2002. "The Case for Incompatibilism." *Philosophy and Phenomenological Research* 64/3: 699–706.
- SMART, J. J. C. 1961. "Free-Will, Praise and Blame." *Mind* 70/279: 291–306.
- SMILANSKY, SAUL. 2000. *Free Will and Illusion*. New York: Oxford University Press.
- STRAWSON, GALEN. 1986. *Freedom and Belief*. New York: Oxford University Press.
- STRAWSON, GALEN. 1994. "The Impossibility of Moral Responsibility." *Philosophical Studies* 75: 5–24.
- STRAWSON, P. F. 1974. "Freedom and Resentment." In *Freedom and Resentment and Other Essays*, 1–28. New York: Routledge.
- TAYLOR, RICHARD. 1992. *Metaphysics*. 4th edn. Englewood Cliffs: Prentice-Hall.
- VAN INWAGEN, PETER. 1983. *An Essay on Free Will*. New York: Oxford University Press.
- WALLACE, R. JAY. 1994. *Responsibility and the Moral Sentiments*. Cambridge, MA: Harvard University Press.
- WATSON, GARY. 1975. "Free Agency." *Journal of Philosophy* 72: 205–20.
- WATSON, GARY. 2003. *Free Will*. 2nd edn. New York: Oxford University Press.
- WIDERKER, DAVID. 1995. "Libertarianism and Frankfurt's Attack on the Principle of Alternative Possibilities." *Philosophical Review* 104/2: 247–61.
- WIDERKER, DAVID, and MICHAEL MCKENNA. 2003. *Moral Responsibility and Alternative Possibilities: Essays on the Importance of Alternative Possibilities*. Burlington: Ashgate.
- WOLF, SUSAN R. 1990. *Freedom within Reason*. New York: Oxford University Press.
- WYMA, KEITH. 1997. "Moral Responsibility and the Leeway for Action." *American Philosophical Quarterly* 34: 57–70.