

believe determinism to be incompatible with free will *merely* because determinism implied them. Many people probably have confused determinism with constraint or control or fatalism or mechanism, and so thought determinism to be incompatible with free will for the wrong reasons.

But if these are bad reasons for thinking free will and determinism are incompatible, there may nonetheless be some good reasons. We may still wonder whether determinism *itself* might not conflict with free will—not because it implies constraint, control, and so on, but *just because it is determinism*. For it seems that if determinism is true, there is only one possible future (hence no garden of many forking paths into the future); and this fact alone seems to rule out the possibility of free will and responsibility for actions.

To this objection, compatibilists issue a challenge of their own. “If there is an argument to show that determinism *must* be incompatible with free will, *just because* it is determinism, and *not* because it implies constraint or control by others or fatalism or mechanism, then provide us with such a direct argument for the incompatibility of free will and determinism! In short, “prove it.” In the next chapter, we will consider how incompatibilists try to meet this challenge.

An Addendum on the Term Soft Determinism

In many writings on free will, compatibilists are often referred to as *soft determinists*. Soft determinists are compatibilists who also believe that determinism is true. Classical compatibilists, such as Hobbes, Hume, and Mill, were also soft determinists, since they believed that determinism was true in addition to believing that freedom and determinism were compatible.

Suggested Reading

A lively and readable defense of compatibilism is Daniel Dennett's *Elbow Room: The Varieties of Free Will Worth Wanting* (MIT, 1984). Defenses of classical compatibilism appear in essays by J.J.C. Smart (in Gary Watson, ed., *Free Will* [Oxford: Oxford University Press, 2nd ed., 2003]) and Kai Nielsen (in Robert Kane, ed., *Free Will*). Other selections from classical compatibilists are contained in Derk Pereboom, ed., *Free Will* (Hackett, 1997); and classical compatibilist positions are discussed in Ilham Dilman's historical introduction, *Free Will* (Routledge, 1999).



Incompatibilism

1. The Consequence Argument

The popularity of compatibilism among modern philosophers and scientists means that *incompatibilists*—those who hold the traditional belief that free will and determinism are in conflict—must provide arguments to support their position. Incompatibilists cannot merely rely on their intuitions about forking paths into the future to make their case, as in chapter 1. They must back up their intuitions with arguments that show why free will and determinism must be incompatible. New arguments for incompatibilism have indeed been proposed in modern philosophy to meet this challenge. The most widely discussed of these new arguments for the incompatibility of free will and determinism is the subject of this chapter.

The argument is called the Consequence Argument, and it is stated informally as follows by one of its proponents, Peter van Inwagen:

If determinism is true, then our acts are the consequences of the laws of nature and events in the remote past. But it is not up to us what went on before we were born; and neither is it up to us what the laws of nature are. Therefore the consequences of these things (including our own acts) are not up to us.¹

To say it is not “up to us” what “went on before we were born,” or “what the laws of nature are,” is to say that there is nothing we can now do to change the past or alter the laws of nature (such things are beyond our control). This gives us two premises of the Consequence Argument.

- (1) There is nothing we can now do to change the past.
- (2) There is nothing we can now do to change the laws of nature.

Putting these two premises together, we get

- (3) There is nothing we can now do to change the past and the laws of nature.

But if determinism is true, then

- (4) Our present actions are the necessary consequences of the past and the laws of nature. (Or, equivalently, it is necessary that, given the past and the laws of nature, our present actions occur.)

So, if determinism is true, it seems that

- (5) There is nothing we can now do to change the fact that our present actions are the necessary consequences of the past and the laws of nature.

But if there is nothing we can now do to change the past and the laws of nature (which is step 3) *and* nothing we can now do to change the fact that our present actions are the necessary consequences of the past and the laws of nature (step 5), it would seem to follow that, if determinism is true (step 4), then

- (6) There is nothing we can now do to change the fact that our present actions occur.

In other words, we *cannot now do otherwise* than we actually do. Since this argument can be applied to any agents and actions at any times, we can infer from it that *if determinism is true, no one can ever do otherwise*; and if free will requires the power to do otherwise, then no one has free will.

2. Assessing the Argument

Van Inwagen thinks the first two premises of this Consequence Argument are undeniable. We cannot now change the past (1) or the laws of nature (2). Step 3 states what appears to be a simple consequence of premises 1 and 2: if you can't change the past or the laws, then you can't change the conjunction of both of them. Premise 4 simply states what is implied by the definition of determinism: if determinism is true, then our actions are the necessary consequences of the past and laws of nature in the sense that they *must* occur, *given* the past and the laws. By asserting premise 4, of course, the argument is assuming the truth of determinism. But it is doing so only hypothetically, in order to show that, *if* determinism is true (premise 4), *then* no one could have done otherwise (6). So the

Consequence Argument does not depend on determinism's actually being true; rather, it seeks to show what determinism would imply (no free will), *if* it were true.

We are left to assess steps 5 and 6. How are they arrived at? Step 5 ("There is nothing we can now do to change the fact that our present actions are the necessary consequences of the past and the laws of nature") follows from premise 4 by virtue of a rule that van Inwagen calls

Rule Alpha. There is nothing anyone can do to change what *must* be the case (or what is necessarily so).

This rule gets us from premise 4 to step 5 in the following way. According to premise 4, it *must be that*, given laws of nature and the past, our present actions occur. But Rule Alpha says no one can now change *what must be*. So it follows that we cannot now change the fact that, given the laws of nature and the past, our present actions occur—which is what step 5 says.

Van Inwagen thinks this Rule Alpha is also undeniable. How, he asks, could anyone change what is necessarily so? If it is necessarily so that $2 + 2 = 4$, then no one can change that; and if someone could change the fact that $2 + 2 = 4$, then it would not be necessarily so.

This brings us to the conclusion of the argument, step (6): "There is nothing we can now do to change the fact that our present actions occur." This conclusion follows from earlier steps, as noted, by virtue of the following inference: if there is nothing we can now do to change the past and the laws of nature (step 3) and nothing we can now do to change the fact that our present actions are the necessary consequences of the past and the laws of nature (step 5), then there is nothing we can now do to change the fact that our present actions occur (6). This inference involves a second rule that van Inwagen calls

Rule Beta. If there is nothing anyone can do to change X, and nothing anyone can do to change the fact that Y is a necessary consequence of X, then there is nothing anyone can do to change Y either.

Rule Beta has been called a "Transfer of Powerlessness Principle." For it says in effect that if we are "powerless" to change X, and if Y is necessarily going to occur if X does, and we are powerless to change that also, then we are also powerless to change Y. In other words, our powerlessness to change X "transfers" to anything that necessarily follows from X.

This Rule Beta also seems intuitively correct, according to van Inwagen. If we can't do anything to prevent X from occurring and Y is *necessarily* going to occur if X does, how could we do anything to prevent Y from occurring? Consider an example. Suppose the sun is going to explode in the year 2050 and there is nothing we can now do to change that fact. There

is also nothing we can now do to change the fact that, if the sun explodes in 2050, all life on earth will end in 2050. If both these claims are true, it seems obvious that there is nothing anyone can now do to change the fact that all life on earth will end in 2050. Here is another example. If there is nothing anyone can now do to change the laws of nature, and nothing anyone can now do to change the fact that the laws of nature entail that nothing goes faster than the speed of light, then there is nothing anyone can now do to change the fact that nothing goes faster than the speed of light.

One could go on adding examples like these supporting Rule Beta. Suffice it to say that Rule Beta does *seem* to be as undeniable as Rule Alpha (which says that no one can change what is necessarily so); and if Rule Beta is also valid, since the other premises of the Consequence Argument seem undeniable, the argument would be both valid and sound, as van Inwagen and other incompatibilists claim. The Consequence Argument would show that determinism conflicts with anyone's power to do otherwise and thus conflicts with free will.

3. An Objection Concerning "Can" and "Power"

The Consequence Argument is a powerful argument for the incompatibility of free will and determinism, and it has swayed many persons. But it is also a controversial argument and has generated much debate. As you would expect, compatibilists and soft determinists reject the Consequence Argument. They must reject it or their views would be refuted in one fell swoop. But where do compatibilists and other critics of the Consequence Argument think it goes wrong, if it goes wrong at all? Most critics of the argument tend to focus on the crucial expression "There is nothing we can now do to change . . ." which appears in many steps of the version of the Consequence Argument presented in section 2. This expression contains the word "can"—one of the most difficult words in the language to interpret.

Talking about what persons "can" (and "cannot") do is talking about their *powers* or *abilities*. So how you interpret persons' powers and abilities has an obvious bearing on the Consequence Argument. For example, compatibilist critics of the Consequence Argument often argue that if you interpret terms like "can," "power," and "ability" in the *hypothetical* way proposed by classical compatibilists, the Consequence Argument will fail. As we saw in chapter 2, according to classical compatibilists, to say

"You *can* (or you have the *power* or the *ability*) to do something"

means there are no *constraints* or *impediments* preventing you from doing it, so that

"You *would* do it, if you chose or wanted to do it."

Such an analysis of "can," "power," or "ability" is called "hypothetical" (or "conditional") because it has an "if" in it. But how does such an analysis refute the Consequence Argument? First, consider the initial two premises of the Consequence Argument: "There is nothing we can now do to change the past" and "There is nothing we can now do to change the laws of nature." On the hypothetical analysis of "can," to say we can change the past or the laws would mean that

"We *would* change the past or the laws of nature, if we chose or wanted to."

Now this claim is false. No persons would change the past or the laws of nature, *even if* they chose or wanted to, because no one has the power or ability to do it. So the initial *premises* of the Consequence Argument come out *true* on this compatibilist analysis. There is nothing anyone can now do to change the past and the laws of nature *even on the hypothetical analysis of "can"* favored by many compatibilists.

But the hypothetical analysis gives a different answer when we consider the *conclusion* of the Consequence Argument: "There is nothing any persons can do to change the fact that their present actions occur," or in other words, "No persons can do otherwise than they actually do." To show why this conclusion fails on the hypothetical analysis of "can," consider a simple everyday action, such as Molly's raising her hand. To say that Molly could have done otherwise than raise her hand (to say, for example, that she could have kept her hand by her side) means, on the hypothetical analysis, that

"She would have done otherwise than raise her hand, if she had chosen or wanted to do otherwise."

Now, as noted in chapter 2, this hypothetical claim can be true even if Molly's action was determined. For the hypothetical claim simply implies that Molly would have done otherwise, *if the past had been different in some way*—that is, if (contrary to fact) she had chosen or wanted differently.

Note that making this hypothetical claim does not imply that Molly could have *changed* the past or the laws of nature from what they actually were. The hypothetical claim merely means that no constraints or impediments would have prevented her from acting differently, *if she had chosen or wanted differently*; and this may well be true even though she did *not* in

fact choose or want differently. In other words, with ordinary everyday actions, such as raising one's hand or getting on a bus, there may *sometimes* be constraints preventing us from doing them or doing otherwise (we may be tied up, paralyzed, or coerced). But often there may be no such constraints preventing us from doing these everyday things; and so we could have done them if we had wanted. By contrast, there are *always* constraints preventing us from changing the past and laws of nature.

As a result, the *premises* of the Consequence Argument come out *true* on the compatibilist hypothetical analysis of "can": Molly *cannot* change the past or the laws of nature, even if she wants to. But the *conclusion* of the Consequence Argument comes out *false*: Molly *can* nonetheless sometimes do otherwise than she actually does (e.g., do otherwise than raise her hand), in the hypothetical sense, because nothing *would* have prevented her, if she had wanted to. So, on the hypothetical analysis, the Consequence Argument would have true premises but a false conclusion, and it would be an invalid argument.

You might wonder at this point what *part* of the Consequence Argument goes wrong in this case—which premise or rule. The answer is Rule Beta. Even defenders of the Consequence Argument, such as van Inwagen, concede that Rule Beta is the hardest part of the argument to defend (though they themselves believe Rule Beta is valid). Rule Beta licenses the inference that gets one to the conclusion of the Consequence Argument (step 6), from steps 1 to 5: if there is nothing we can now do to change the past and the laws and nothing we can now do to change the fact that our present actions are the necessary consequences of the past and the laws, then we cannot now do otherwise than we actually do. On the compatibilist hypothetical analysis of "can," the premises of this inference are true, while its conclusion is false. For on the hypothetical analysis of "can" there *is* nothing we can now do to change the past and the laws of nature, but there is something we can now do to change ordinary actions, such as raising our hand. Rule Beta is therefore invalid (it has counterexamples); and the Consequence Argument fails.

4. Defenders of the Consequence Argument Respond

Now this objection to the Consequence Argument works, of course, only if the hypothetical analysis of "can," "power," or "ability" favored by classical compatibilists is correct. But why should we believe this hypothetical analysis of "can" and "power"? Defenders of the Consequence Argument, such as van Inwagen and Carl Ginet, see no good reason to believe in the compatibilists' analysis of these notions and so they typically

respond to the above argument in the following way:

So the hypothetical analyses of "can" (or "power" and "could have done otherwise") that you compatibilists favor would refute Rule Beta and the Consequence Argument. Should that make us incompatibilist defenders of the Consequence Argument doubt Rule Beta and the Consequence Argument? Not at all. It just gives us another reason for doubting your compatibilist hypothetical analysis of "can," which we never thought was very plausible in the first place. If your analysis allows you to say that Molly can do otherwise (than raise her hand), even though she can't change the past and the laws of nature and even though her action (of raising her hand) is a necessary consequence of the past and the laws of nature, *then something must be wrong with the hypothetical analysis* of "can" that you compatibilists favor. The premises and rules of the Consequence Argument, including Rule Beta, seem more intuitively true to us than any hypothetical analysis of "can." So, if we have to reject one or the other, we would reject your compatibilist analysis rather than the Consequence Argument. In fact, hypothetical analyses of "can" and "could have done otherwise" that many compatibilists favor are subject to serious objections anyway. So they should be rejected in any case and not just because one favors the Consequence Argument.²

What are the "serious objections" to hypothetical analyses of "can" and "could have done otherwise" referred to in this passage? The objection that many philosophers regard as the most serious goes like this: hypothetical analyses of "can" and "could have done otherwise" sometimes (wrongly) tell us that agents can do otherwise, or could have done otherwise, in cases where it is clear that the agents could *not* have done otherwise. So the hypothetical analyses must be wrong. Here is an example of Michael McKenna's illustrating this objection. Suppose that Danielle has been scarred by a terrible childhood accident involving a blond Labrador retriever. The accident rendered her

psychologically incapable of wanting to touch a blond haired dog. Imagine that, on her sixteenth birthday, unaware of her condition, her father brings her two puppies to choose between, one being a blond haired Lab, the other a black haired Lab. He tells Danielle just to pick up whichever of the two she pleases and that he will return the other puppy to the pet store. Danielle happily, and unencumbered, does what she wants and picks up the black Lab.³

Was Danielle free to *do otherwise* (*could* she have done otherwise) than pick up the black Lab? It seems not, McKenna says. Given her traumatic childhood experience, she cannot even form a *want* to touch a blond-haired Lab, hence she could not pick up one.

But notice that the compatibilist hypothetical analysis of “she could have done otherwise” would be true in this case: *If* Danielle *did* want to pick up the blond-haired Lab, then she would have done so. So the hypothetical analysis gives us the wrong answer in this case and in many other similar cases. It tells us Danielle could have done otherwise (because she would have, if she had wanted), when in fact she could *not* have done otherwise (because she could not have *wanted* to do otherwise).

The problem with the hypothetical analysis brought out by this example is the following: to truly capture the meaning of “She *could* have done otherwise,” it is not good enough to simply say “She *would* have done otherwise, *if* she had wanted to”; one must add “*and she could also have wanted* to do otherwise.” But then the hypothetical analysis merely pushes the question of whether the agent could have *done* otherwise back to another question of whether the agent could have *wanted* or *chosen* (or *willed*) to do otherwise. And answering this further question requires another “could” statement (“She could have wanted or chosen to do otherwise”), which in turn requires another hypothetical analysis: “She would have wanted or chosen to do otherwise, *if* she had *wanted* or *chosen* to want or choose otherwise.” And the same question would arise about this further hypothetical analysis, requiring yet another “could” statement to be analyzed, and so on indefinitely.

The result is an infinite regress that would never allow one to eliminate the word “could” and would never allow one to definitively answer the original question of whether the agent could have done otherwise—which shows that something has gone wrong with the hypothetical analysis. For reasons such as this, defenders of the Consequence Argument think the hypothetical analysis of “could have done otherwise” favored by classical compatibilists is flawed. Such an analysis would undermine the Consequence Argument, if it were correct. But there are reasons to think it is not correct.

At this point, debates about the Consequence Argument tend to reach an impasse. Defenders of the Consequence Argument think its premises and rules are far more plausible than any compatibilist analysis of “could have done otherwise” (hypothetical or otherwise), while compatibilists obviously think the opposite. Many compatibilists today do concede that the *classical* compatibilist analysis of “could have done otherwise” may be flawed, for the reasons just given or for other reasons. But these same modern compatibilists insist that defenders of the Consequence Argument are begging the question when they assume that *no* compatibilist analysis of “could have done otherwise” could possibly be right, merely because the classical compatibilist analysis is flawed.

Perhaps this is so. But then the burden of proof lies with compatibilists to give a better account of “could have done otherwise” than classical compatibilists have offered—or to find some other way to refute the Consequence Argument. We shall see in later chapters that modern compatibilists have tried to do one or another of these two things. Some modern compatibilists have sought better compatibilist analyses of “could have done otherwise.” Others have sought entirely new ways of refuting the Consequence Argument.

Suggested Reading

Van Inwagen’s defense of the Consequence Argument is in his *An Essay on Free Will* (Oxford: Clarendon, 1983). The Consequence Argument is also defended by Carl Ginet in *On Action* (Cambridge, 1990). Other discussions for and against the Consequence Argument are included in the collections of readings cited in the suggested readings of chapter 1.



Libertarianism, Indeterminism, and Chance

1. *Libertarianism Defined*

Even if some argument for incompatibilism, such as the Consequence Argument, should succeed, that success would not by itself show that we have free will. A successful argument for incompatibilism would show only that free will and determinism cannot both be true. If one is true, the other must be false. Thus, incompatibilists may go in either of two directions. They may affirm free will and deny determinism, or affirm determinism and deny free will. Incompatibilists who affirm free will and deny determinism are called *libertarians* in modern free will debates. It is this libertarian view that we are now going to consider. (The opposing view—affirming determinism and denying free will—is called hard determinism, and it will be considered in chapter 7.)

People who are libertarians about free will see themselves as defenders of the “deeper” freedom of the will of chapter 1, which they believe to be incompatible with determinism. This deeper freedom, as libertarians see it, is the “true” free will that most people have traditionally believed in before they began to worry about determinism. From the libertarian point of view, compatibilists give us only a pale image of this true freedom (a “wretched subterfuge,” as Immanuel Kant said); libertarians claim to give us the real thing. But giving us the real thing (if libertarian free will really is the real thing) turns out to be more difficult than one may at first imagine, as we shall see in this chapter and the next.

Libertarianism will thus be defined from this point onward as the view that (1) free will and determinism are incompatible (incompatibilism),

(2) free will exists, and so (3) determinism is false. Libertarianism in this sense—libertarianism *about free will*—should not be confused with the political doctrine of libertarianism, the view that governments should be limited to protecting the liberties of individuals as long as the individuals do not interfere with the liberties of others. Libertarianism about free will and political libertarianism share a name—from the Latin *liber*, meaning “free”—and they share an interest in freedom. But libertarians about free will are not necessarily committed to all the views about limited government held by political libertarians. Libertarians about free will can in fact (and many do) hold different political views—conservative, liberal, libertarian, or whatever—so long as they share a commitment to the ideal of persons having responsibility for their actions and their lives in an ultimate sense that is incompatible with determinism.

2. *The Libertarian Dilemma: Ascent and Descent Problems*

To defend libertarianism about free will, one obviously has to do more than merely argue for the incompatibility of free will and determinism, as important as that may be. One must also show that we can actually have a free will that is incompatible with determinism. Many people believe that an incompatibilist free will of the kind that libertarians affirm is not even possible or intelligible and that it has no place in the modern scientific picture of the world. Critics of libertarianism note that libertarians have often invoked obscure and mysterious forms of agency or causation to defend their view.

To explain how free actions can escape the clutches of physical causes and laws of nature, libertarians have posited transempirical power centers, nonmaterial egos, noumenal selves outside space and time, unmoved movers, uncaused causes, and other unusual forms of agency or causation—thereby inviting charges of obscurity or mystery against their view. Even some of the greatest defenders of libertarianism, such as Immanuel Kant, have argued that we need to believe in libertarian freedom to make sense of morality and true responsibility, but we cannot completely understand such a freedom in theoretical and scientific terms.

The problem that provokes this widespread skepticism about libertarian free will has to do with the dilemma mentioned in chapter 1 and touched upon in chapter 2: if free will is not compatible with determinism, it does not seem to be compatible with *indeterminism* either. Let us call this the “Libertarian Dilemma.”¹ Events that are undetermined, such as quantum

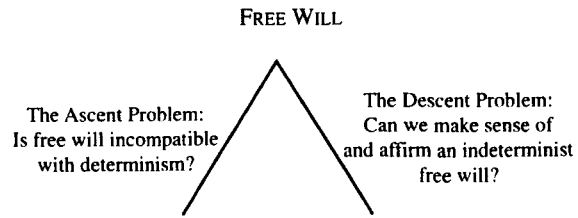


Figure 4.1 Incompatibilist Mountain and the Libertarian Dilemma

jumps in atoms, happen merely by chance. So if free actions must be undetermined, as libertarians claim, it seems that they too would happen by chance. But how can chance events be free and responsible actions? To solve the Libertarian Dilemma, libertarians must not only show that free will is *incompatible* with *determinism*, they must also show how free will can be *compatible* with *indeterminism*.

Imagine that the task for libertarians in solving this dilemma is to ascend to the top of a mountain and get down the other side. (Call the mountain "Incompatibilist Mountain": figure 4.1). Getting to the top consists in showing that free will is incompatible with determinism. (Call it the Ascent Problem.) Getting down the other side (call it the Descent Problem) involves showing how one can make sense of a free will that requires *indeterminism*.

Getting to the top of this mountain—demonstrating that free will and determinism are incompatible—is a difficult enough task for libertarians, as we have seen in chapter 3. But many critics of libertarianism believe the Descent Problem—making sense of a free will that requires indeterminism—is even more difficult. Mountain climbers say that the descent from a mountain peak is often more difficult and dangerous than the ascent; and this may be the case for libertarians. The air is thin and cold up there on Incompatibilist Mountain; and if you stay up for any length of time, say critics of libertarianism, your mind gets foggy. You start having visions of fantastical ideas, such as transempirical power centers, noumenal selves, and unmoved movers, which libertarians have often invoked to explain their view.

3. *Indeterminism the Bogyman*

Why is it so difficult to make sense of a free will that requires indeterminism (and hence to solve the Descent Problem) without slipping into mystery or obscurity? Some of the difficulties that indeterminism poses for free will were suggested in earlier chapters. But let us see if we can get an overview of them.

1. First, one often hears critics of libertarianism argue that events that are undetermined happen merely by chance and are not under the *control* of anything, hence are not under the control of the agent. It is not "up to" agents whether undetermined events occur or not. But if events are not under the control of an agent, they cannot be free and responsible actions.

2. A related argument was suggested in chapter 1. Suppose a choice was the result of a quantum jump or other undetermined event in a person's brain. Would this amount to a free and responsible choice? Such undetermined effects in the brain or body would be unpredictable and impulsive—like the sudden occurrence of a thought or the spasmodic jerking of an arm that one could not have predicted or influenced—quite the opposite of what we take free and responsible actions to be. It seems that undetermined events happening in the brain or the body would occur *spontaneously* and would be more likely to *undermine* our freedom rather than to *enhance* our freedom.

3. Nor would it help to suppose that the indeterminism or chance came *between* our choices and our actions. Imagine that you have chosen to make a delicate cut in a fine piece of cloth, but because of an undetermined twitching in your arm, you make the wrong cut. In this case, the undetermined twitching in your arm was no enhancement of your freedom, but a hindrance or obstacle to your carrying out your intended purposes. Critics of libertarian freedom often contend that this is what indeterminism would always be—a *hindrance* or *impediment* to freedom. It would get in the way, diminishing rather than enhancing *control* and *responsibility* for what happens. Note that the twitching of your arm is actually a *constraint* on your freedom in the classical compatibilist sense, since it *prevents* you from doing what you *want* to do, that is, make the delicate cut properly. So, far from giving us more freedom, it seems that indeterminism would turn out to be another kind of impediment limiting our freedom.

4. Even more absurd consequences follow if we suppose that indeterminism or chance is involved in the initiation of everyday actions. A nineteenth-century critic of undetermined free action, Arthur Schopenhauer, imagined the case of a man who suddenly found his legs start to move *by chance*, carrying him across the room against his wishes.² Is this what libertarians have in mind, Schopenhauer asked, when they insist that free actions must be undetermined? Such caricatures are popular among critics of indeterminist freedom for obvious reasons: undetermined or chance-initiated actions would represent the opposite of free and responsible actions.

5. Going a little deeper, critics of libertarian freedom also note that, if choices or actions are undetermined, they may occur otherwise, *given exactly the same past and laws of nature*. This follows, as we saw, from

indeterminism, which implies different possible futures, given the same past. But such a requirement has troubling consequences regarding free choices, as noted in chapter 2. Here is a further example illustrating the problem. Suppose Mike, who is deliberating about whether to vacation in Hawaii or Colorado, gradually comes to favor and choose Hawaii. If Mike's choice, when he finally makes it, was undetermined, as libertarians require, then he might have chosen otherwise (chosen to visit Colorado instead), given exactly the same deliberation up to the moment of choice that in fact led him to favor and choose Hawaii (the same thoughts, reasoning, beliefs, desires, and so on). As noted in our discussion of Molly's choosing a career, it is difficult to make sense of this. Mike's choosing Colorado in such circumstances (in which he had come to favor Hawaii) would seem irrational and inexplicable, capricious and arbitrary. If the choice of Hawaii came about by virtue of undetermined events in Mike's brain, this would not be an occasion for rejoicing in his freedom, but for consulting a neurologist about the waywardness of his neural processes.

4. *Reasons, Randomness, and Luck*

6. At this point, some defenders of indeterminist freedom appeal to the claim of the eighteenth-century philosopher Gottfried Leibniz, that prior reasons or motives need not determine choice or action, they may merely "incline without necessitating."³ For example, Mike's reasons for wanting to vacation in Colorado (he likes skiing and wants to meet friends there) might "incline" him to choose Colorado over Hawaii. But these reasons do not "necessitate" or determine that he will choose Colorado. Similarly his reasons for favoring Hawaii (he also likes beaches and surfing) incline him toward Hawaii without determining that choice.

Leibniz's claim that reasons may "incline without necessitating" is an important one. But, unfortunately, it will not solve the problem about Mike's choice described in objection 5. For it is precisely *because* Mike's prior reasons and motives (his beliefs and desires about beaches and surfing) inclined him more strongly toward the choice of Hawaii than his choosing Colorado by chance at the end of the same deliberation would be arbitrary, irrational, and inexplicable. Similarly, if his reasons had inclined him more strongly toward Colorado, then choosing Hawaii by chance at the end of the same deliberation would have been irrational and inexplicable.

What if Mike's prior reasons and motives had not inclined him more strongly to *either* alternative? Then, if the choice were undetermined, matters would be even worse. For the choice would then be doubly arbitrary—arbitrary either way he might choose. Medieval philosophers,

who discussed free will, had a name for the condition of an agent who has no better reasons for choosing one option rather than the other. They called it "liberty of indifference." You have probably heard the well-known illustration of the liberty of indifference involving Buridan's ass—the donkey that starved between two equidistant bales of hay because it had no reason to choose one over the other.

Jean Buridan was a medieval French philosopher to whom this famous example of the donkey is often wrongly attributed. The original example goes back to the medieval Arabic philosopher Al-Ghazzali, who imagined a camel starving between two groves of date trees. These examples of the liberty of indifference were often used later by philosophers, such as Hume and Schopenhauer, to ridicule libertarian or indeterminist free will. (Al-Ghazzali had used his example for a similar purpose.) Of course, a human, who was not an ass, would undoubtedly not starve to death in these conditions. It would be better to flip a coin and choose one option arbitrarily or by chance than to go without food altogether. But such a solution to the liberty of indifference—choosing by a coin flip—still amounts to choosing arbitrarily or by chance. Is that what indeterminist freedom amounts to?

7. Indeed, another frequently heard objection to indeterminist free will is precisely that undetermined free choices must *always* amount to mere *random* choices, like flipping a coin or spinning a wheel to select from among a set of alternatives. Perhaps there is a role for random choices in our lives—for sometimes settling choices by a coin flip or spinning a wheel—when we are indifferent to the outcomes. (Which movie should I see tonight when I like both available options?) But suppose that *all* our free and responsible choices—including momentous ones, like whether to act heroically or treacherously, to lie to a friend, or to marry one person rather than another—had to be settled by random selection in this way. Such a consequence, according to most philosophers, would be a reduction to absurdity of the view that free will and responsibility require indeterminism.

8. Finally, consider the following objection, which has been suggested by a number of critics of indeterminist free choice.⁴ We may call it the "Luck Objection." Indeterminism, as noted earlier, implies different possible futures, given exactly the same past. Suppose then that two agents had exactly the same pasts up to a point at which they were faced with a choice between distorting the truth for selfish gain or telling the truth at great personal cost. One agent lies and the other tells the truth. Bruce Waller summarizes this objection as follows: if the pasts of these two agents "are really identical" in every way up to the moment of choice, "and the difference in their acts results from chance," would there "be any

grounds for distinguishing between [them], for saying that one person deserves censure for a selfish decision and the other deserves praise?"⁵

Another critic, Alfred Mele, poses the same problem in terms of a single agent in different possible worlds. Suppose that in the actual world, John fails to resist the temptation to do what he thinks he should not do, arrive on time at a meeting. If John could have done otherwise given the same past, then we could imagine that his counterpart, John*, in an alternative possible world (which is exactly the same as the actual world up to the moment of choice) resists the temptation and arrives on time. Mele then argues that "if there is nothing about the agents' powers, capacities, states of mind, moral character and the like that explains this difference in outcome, . . . the difference is just a matter of luck." It would seem that John* got lucky in his attempt to overcome temptation, while John did not. Would it be fair or just to reward the one and punish the other for what appears to be ultimately the luck of the draw?⁶

5. *The Indeterminist Condition and Extra Factor Strategies*

Objections such as the eight outlined in sections 3 and 4 lie behind the many charges often heard in the history of free will debates against libertarian free will—charges that undetermined actions would be "arbitrary," "capricious," "random," "uncontrolled," "irrational," "inexplicable," or "matters of luck or chance"—anything but free and responsible actions. The first task for libertarians, if they are to make sense of their view and solve the Descent Problem, is to address these familiar charges.

To understand how libertarians have gone about the task of trying to answer these charges, it helps to note that the problem lying behind all the objections just given is the problem of reconciling free actions with what we may call

The Indeterminist Condition: the agent should be able to act and act otherwise (choose different possible futures), *given the same past circumstances and laws of nature.*

It is this Indeterminist Condition that makes it seem irrational and inexplicable, capricious and arbitrary, for Mike to choose to vacation in Colorado given the same prior deliberation that in fact led him to favor and choose Hawaii. It is the same Indeterminist Condition that leads Mele to argue that if the circumstances of John and John* are exactly the same up to the moment of choice (if there is no difference in their "powers, capacities, states of mind, moral character and the like"), then "there is

nothing about the agents that explains" why John failed to overcome the temptation and John* did not—except luck.

Reflecting on this Indeterminist Condition gives us some insight into the strategies libertarians have traditionally employed in their attempts to make sense of libertarian free will. Libertarians have typically reasoned in the following way. If agents may act or act otherwise, given the same past circumstances and laws of nature, then some *additional* factor *not included among the past circumstances or laws* must account for the difference in outcome—for an agent's acting or choosing in one way rather than the other. The agent's acting differently cannot be accounted for solely by the circumstances of the agent prior to action because, by hypothesis, there is no difference in these prior circumstances. So if the outcome is not to be merely random, arbitrary, and inexplicable, an extra factor must be involved over and above the past circumstances and laws to account for it.

Let us call any such strategy for making sense of libertarian free will an "extra-factor strategy." Throughout history, libertarians have regularly invoked some extra factor or other to explain how free will is possible in their sense. But the extra factors have varied. Libertarians have invoked immaterial minds or souls, noumenal selves outside space and time, special forms of agent causation that cannot be reduced to scientific modes of causation, "acts of will" or "volitions" that cannot by nature be determined by prior events, "reasons" or "purposes" or "final causes" that explain actions without being antecedent causes of actions, and so on. These extra factors are meant to explain why free choices or actions do not merely occur in an arbitrary, capricious, random, uncontrolled, or irrational way—even though the choices or actions are undetermined by prior causes and laws.

In the next chapter, we shall consider some of the most important traditional extra-factor strategies by which libertarians have attempted to make sense of the deeper kind of free will they believe in.

Suggested Reading

There are many critiques of the libertarian position on free will. Three readable critiques are Richard Double, *The Non-reality of Free Will* (Oxford, 1991), Bruce Waller, *Freedom Without Responsibility* (Temple, 1990), and Ted Honderich, *How Free Are You?* (Oxford, 1993). A useful collection of readings for and against libertarian views of freedom is *Agents, Causes, and Events: Essays on Free Will and Indeterminism*, edited by Timothy O'Connor (Oxford, 1995).



Minds, Selves, and Agent Causes

1. Mind-Body Dualism

The most obvious extra-factor strategy that comes to mind when people think about how to make sense of libertarian free will involves a *dualism* of mind and body (such as that of René Descartes.) If the “mind” or “soul” were distinct from the body, it would be outside the physical world and its activity would not be governed by laws of nature that govern physical events. If, in addition, a disembodied mind or soul could interact with the physical world by influencing the brain, as Descartes imagined, then the mind or soul would be the “extra factor” libertarians need to explain free choice. Whatever could not be fully explained by the activity of brain or body might be explained by the activity of the mind or soul.

For such a dualist solution to the free will problem to work, the physical world would have to cooperate, allowing some indeterminism in nature, perhaps in the brain. It may be true that quantum jumps or other undetermined events in the brain would not by themselves amount to free choices. But undetermined events in the brain might provide the “leeway” or “causal gaps” in nature through which an extra factor, such as an immaterial mind or soul, might intervene in the physical world to influence physical events.

Those who take this dualist approach to free will could thus accept the Indeterminist Condition in a qualified form: they could say that free agents are able to choose or choose otherwise, all past *physical* circumstances remaining the same (because physical circumstances are the kind that are governed by laws of nature). But the activity of the agent’s mind or soul would not be among the physical circumstances and would not be governed by laws of nature; and the activity of an immaterial mind or soul

could account for why one choice was made rather than another. Thus free choices would not be arbitrary, random, or inexplicable after all; nor would they occur merely by chance or luck, even though it might look that way, if one just described the physical world.

This dualist solution to the free will problem has been tempting through the ages and still is. Many people naturally tend to think mind-body dualism is the obvious and perhaps the only way to solve the free will problem. So it is important to understand why many philosophers believe that affirming a dualism of mind and body will not by itself solve the problems about libertarian free will discussed in chapter 4. Let us put aside for the moment the usual philosophical concerns people have about an “interactionist” mind-body dualism of the kind posited by Descartes: How does an immaterial mind act on a physical body? Where does the mind act on the body? Are the laws of nature violated by the intervention of the mind, and if so, how? Whatever problems of these kinds a dualism of mind and body may have, the point of interest for us is that an appeal to mind-body dualism will not of itself solve the problems about free will posed by indeterminism that we have been considering.

To see why, ask the following question: if a free choice (such as Molly’s choice to join the law firm in Dallas or Mike’s to vacation in Hawaii or John’s to arrive late) is not determined by the prior *physical* activity of the agent’s brain, is the choice determined by the prior *mental* activity of the agent’s mind or soul? Dualists who are libertarians about free will must answer that free choices in a libertarian sense cannot be determined by the prior activity of a disembodied mind or soul any more than free choices can be determined by prior physical activity of the body. For, determinism either way would rule out the possibility of doing otherwise, hence would rule out libertarian free will. If God had so made us that the activities and effects of our *minds* were also determined, we would be no better off *regarding free will* just because our minds were separate from our bodies.

But if determinism by the mind is no more acceptable than determinism by the body, then dualists who want to defend libertarian free will cannot merely say that Molly (or Mike or John) could have chosen or chosen otherwise, given all the same past *physical* circumstances. Dualists must also say that free agents could have chosen or chosen otherwise, given all the same past *physical and mental* circumstances. If dualists do *not* say this, they will not really have avoided determinism. But if dualists *do* say this, all the original problems about the Indeterminist Condition will come back to haunt them. If Molly might have chosen the law firm in Austin, given all the same prior thoughts, reasoning, and other mental (as well as physical) circumstances that in fact led her to favor the Dallas firm, then

her choice to join the Austin firm would have been just as irrational, inexplicable, and arbitrary if it issued from a disembodied mind or soul as it would if it had issued from an embodied person. If John and John* might have chosen differently, given exactly the same mental (and physical) histories up to the moment when they did choose, then Mele's question comes back to haunt us: "What can account for the difference in their choices—why John failed to overcome the temptation and John* did not—except luck?"

For reasons such as these, placing the agent's thoughts and deliberations in a disembodied mind or soul does not solve the problems about an undetermined free will. Dualism simply transfers these problems to another level, from the physical sphere to the mental. That is why a critic of libertarianism, such as Simon Blackburn, can say: "The dualist approach to free will makes a fundamental philosophical mistake. It sees a problem and tries to solve it by throwing another kind of 'thing' into the arena [the controlling soul]. But it forgets to ask how the new 'thing' escapes the problems that beset ordinary things. . . . If we cannot understand how human beings are free [in a libertarian sense], we cannot understand how [a disembodied mind] can be free" either.¹ Of course, Blackburn's comment does not mean that dualism is necessarily false. But it does mean that appealing to a mind or soul separate from the body will not by itself solve the problem of free will, as some people have believed.

Dualists might appeal to mystery at this point. "We don't know very much about disembodied minds or soul-substances or how they operate," they may say. "How can we be sure an immaterial mind could not make undetermined choices that are not merely random, arbitrary, capricious, and inexplicable?" True enough. We do not know. But if dualists rely on this response and do nothing more, they merely confirm the most common criticism made of libertarian theories of free will—that one cannot make sense of libertarian free will without ultimately appealing to mystery of some kind or other. A great twentieth-century physicist, Erwin Schrödinger, once said something relevant to this point: "At the price of mystery," he said, "you can have anything"—though, we might add, in the words of Bertrand Russell, that you get it too easily, acquiring it by theft rather than honest toil.

2. Kant and Noumenal Selves

Some libertarians concede that libertarian free will is, and must always remain, mysterious. As noted earlier, Immanuel Kant thought libertarian freedom was necessary to make sense of morality and true responsibility.

But Kant also held that a libertarian freedom could not be understood in theoretical or scientific terms.² Science and reason, said Kant, can tell us only the way things *appear* to us in space and time—the world of *phenomena*. But science and reason cannot tell us about the way things are in themselves—the *noumena*. Thus, when scientists try to explain why an agent makes one free choice rather than another, if they are biochemists or neurologists, they will appeal to prior states and processes of agent's brain and body, which appear to us in space and time. If the scientists are psychologists, they will appeal to prior states and processes of the agent's mind which, according to Kant, appear to us in time, but not space. But, in either case, the scientists will fail to explain why one free choice occurs rather than another. For, if the choices are undetermined, it seems that the occurrence of one free choice rather than another cannot be adequately explained by prior states and processes of any kinds, physical or mental.

Now Kant in fact believed that all events occurring in space and time were determined. Writing in the eighteenth century, Kant was convinced that the mechanistic physics of Newton provided the true explanation of the physical world and that this physics was deterministic. But we do not have to assume that science is deterministic, as Kant did, to arrive at a conclusion like his—that free choices cannot be explained by science. For *viewed from science's perspective within space and time*, if free choices were *not* determined, then they would appear to be merely random events, such as quantum jumps in atoms. Either way—determined or random—they would not be free choices. So, had Kant known modern physics, he might have responded in this way: "Free choices can no more be explained by an indeterministic (quantum) physics than they can be explained by a deterministic (Newtonian) physics. I may have been wrong about the truth of Newton's physics. But I was not wrong in concluding that free choices are beyond scientific explanation."

Yet, we also know Kant thought we had to believe in libertarian free will even if science could not explain it. Such a free will was presupposed by our *practical* reason, and, in particular, by our moral life.³ When we deliberate in practical life about whether to keep a promise to a friend, Kant reasoned, we must presuppose we can keep the promise *or* break it and that it is "up to us" what we do. If we did not believe this, deliberating would make no sense. But if we can keep the promise or break it, then the law governing our behavior is a moral law ("You ought to keep your promises") that we can choose to follow *or* to violate.

Kant believed that being governed by such a moral "law" is quite different from being governed by scientific "laws" of nature. Laws of nature are imposed upon us from outside and we cannot choose whether or not to obey them. By contrast, to act in accordance with a moral law is to

be, in Kant's terms, *self-legislating* or *autonomous* (from the Greek *auto* [self] and *nomos* [law]). It is to be governed by a law we give to ourselves, a law we can choose to obey or not obey. Kant held that in our practical moral lives, we must suppose ourselves to be self-legislating or autonomous beings. Such *autonomy*—which amounted to *free will* for him—is not compatible with being governed by scientific laws of nature.

As a result, there is a difference (and a tension) in Kant's view between our *practical* or moral reasoning, which requires that we believe in libertarian free will, and our *theoretical* or scientific reasoning, which cannot explain this freedom. Kant tried to lessen this tension by claiming that science and reason describe the self only as it appears to us in space and time (the phenomenal self), not the self or person as it is "in itself" (the noumenal self). Our real or noumenal selves can be free, he argues, because they are not subject to the constraints of space and time or the laws of nature.

But when science and reason try to explain *how* the noumenal self can be free, they inevitably look for physical, psychological, or social causes of our behavior; and then the scientists are describing only the self as it appears to us, the phenomenal self, not the noumenal or real self. Indeed, anything we might say *about* this noumenal self—about its states or activities—would be describing its physical, psychological, or social circumstances, hence would be describing the phenomenal, not the real, self. The noumenal self is thus the "extra factor" in Kant's theory that is supposed to account for free will. But we cannot say *how* it does so. If free will were the product of a noumenal self in Kant's sense, it would indeed be a mystery.

3. Agent-causation

You can see from the preceding discussion why many modern philosophers who would like to believe in libertarian free will are not satisfied with either mind-body dualist or Kantian solutions to the free will problem. Both dualist and Kantian views require strong and controversial metaphysical assumptions without at the same time solving the problems about indeterminism and chance that make most people reject libertarian free will in the first place. The third traditional libertarian strategy we are going to consider has been more popular among contemporary philosophers. Sometimes this third strategy is combined with other libertarian strategies, such as dualism; but more often it is defended on its own.

This third libertarian strategy is often called an *agent-causal* strategy—or a theory of *agent-causation*—because it focuses on the notion of causation by agents. Free agents are capable of causing their own free acts

in a special way, according to agent-causal views, a way that is not reducible to causation by circumstances, events, or states of affairs. Here is how Roderick Chisholm, a well-known defender of this kind of view, puts the matter:

If we consider only inanimate natural objects, we may say that causation, if it occurs, is a relation between *events* or *states of affairs*. The dam's breaking was an event that was caused by a set of other events—the dam being weak, the flood being strong, and so on. But if a man is responsible for a particular deed, then . . . there is some event [his deed or action] . . . that is caused, *not* by other events or states of affairs, but by the agent, whatever he may be.⁴

Chisholm is suggesting a way out of the Libertarian Dilemma: libertarian free actions cannot be completely *caused* by prior circumstances, events, or states of affairs; and neither can they be *uncaused* or happen merely by chance. But there is a third possibility: we can say that free actions are indeed caused, but not by prior circumstances, events, or states of affairs. Free actions are caused by the *agent* or *self*, which is not a circumstance, event, or state of affairs at all, but a *thing* or *substance* with a continuing existence. We do not have to choose between determinism by prior causes or indeterminism or chance. We can say that free actions are *self-determined* or *agent-caused* even though they are undetermined by events.

Thus the "extra factor" that explains free will for agent-causalists is the agent. Or, to be more precise, the extra factor is a special or unique kind of causal *relation* between an agent and an action that is not reducible to, and cannot be fully explained in terms of, the usual kinds of causation by events, occurrences, and states of affairs, either physical *or* mental. The Indeterminist Condition can thus be true in a general sense on the agent-causal view: the agent may act or act otherwise, given all the same past physical *and* mental circumstances and laws of nature because the factor that makes the difference is causation by something (the agent) that is not a *circumstance* at all in the sense of an event or occurrence or state of affairs, whether physical or mental.

Agent-causation of such a *non-event* or *non-occurrent* kind is unusual, as even its defenders, such as Chisholm, acknowledge. (To indicate its special nature, the expression "agent-causation" is often hyphenated in writings on free will, a practice I am following.) We do in fact regularly speak of *things* or *substances* causing events or occurrences: "The stone broke the window." "The cat caused the lamp to fall." But causation by things or substances can usually be interpreted in everyday life as the causation of events or occurrences by other events or occurrences. It is the

stone's *moving* and *striking* the window that caused it to break; and it is the cat's *leaping* onto the table and *hitting* the lamp that caused it to fall. These are *events* involving the stone and the cat, respectively.

But no such paraphrasing in terms of events or occurrences is possible in the case of agent-causation of the non-event or non-occurrent kind that is supposed to explain free will. Agents *non-occurrently* cause things to happen, not by virtue of doing something else or as a result of being in certain states or undergoing changes. In order to account for free actions that are undetermined by prior circumstances, agent-causalists argue that we must recognize another kind of causation alongside the usual causation of events or occurrences by other events or occurrences recognized by the sciences. We must recognize the possibility of direct causation of an event or occurrence by an agent or substance that is a primitive relation, not further analyzable into causation by events or occurrences.

Chisholm illustrates this idea of direct agent-causation by reference to a quotation from Aristotle's *Physics*: "A staff moves a stone, which is moved by a hand, which is moved by a man."⁵ The staff's moving the stone is an instance of ordinary causation of an event by another event, which Chisholm calls *transeunt causation*: it is the staff's *moving* that moves the stone. Similarly, the hand's *moving* causes the staff to *move*, so the hand's moving the staff is another instance of transeunt or event causation. But what are we to say of the movement of the hand by the agent? Chisholm answers as follows:

We *may* say that the hand was moved by the man, but we may *also* say that the motion of the hand was caused by the motion of certain muscles; and we may say that the motion of the muscles was caused by certain events that took place within the brain. But some event, and presumably one of those that took place within the brain, was caused by the agent and not by any other events.⁶

In other words, if we are going to say finally that the *agent* did anything for which the agent was responsible, then sooner or later we must say that the agent *directly* caused some event or other in this chain of events (say an event in the brain or a choice to move the stone), not *by* doing something else and *not* by being caused to do it by any other events. As another agent-cause theorist, Richard Taylor, has put it, "some . . . causal chains . . . have beginnings, and they begin with the agents themselves."⁷

Chisholm calls this direct causation by an agent *immanent causation*, to distinguish it from transeunt causation. He adds:

If what I have been trying to say [about immanent causation] is true, then we have a prerogative which some would attribute only to God: each of us when

we act is a prime mover unmoved. In doing what we do, we cause certain events to happen, and nothing—or no one—causes us to cause those events to happen.⁸

On what grounds does Chisholm say that the agent's immanently causing an event is not caused by other events? The answer, according to Chisholm and other agent-causalists, is that agents are not themselves events or occurrences; so they are not the *kinds* of things that by their nature can be transeuntly caused by other events. If the agent's immanently causing an action could be explained in terms of other events involving the agent (such as states and processes of the agent's brain or mind), then we could ask what caused those other events, and the causal chain would not begin with the agent. But the distinguishing feature of non-event or non-occurrent agent-causation is that it *cannot* be explained in terms of events or occurrences involving the agent. The agent immanently causes an action or event directly and not *by* doing anything else. So there is no other occurrence or event about which to ask: what caused *it*? The causal chain begins with the agent, who is a "prime mover unmoved."

4. Assessing the Agent-causal View: Reid and Causal Power

What are we to say of this agent-causal view? It is not surprising that many critics of libertarian theories of free will find the notion of immanent causation as mysterious as Kantian noumenal selves or Cartesian immaterial minds. To say, as Chisholm does, that we are "prime movers unmoved" or "uncaused causes," like God, does not help, according to these critics, since it merely attempts to explain the obscure by the more obscure. What do we know of how God moves without being moved? And are we humans really like God in this respect, since we *are* clearly moved, at least in part, by many physical, psychological, and social factors, some of which are beyond our awareness?

Even some defenders of agent-causation admit that the notion is mysterious. Richard Taylor, mentioned earlier, says: "One can hardly affirm such a theory of agency with complete comfort . . . and wholly without embarrassment, for the conception of men and their powers which is involved in it, is strange indeed, if not positively mysterious."⁹ Yet Taylor thinks such a notion of agent-causation is the only one consistent with libertarian free agency. "If I believe that something not identical to myself was the cause of my behavior—some event wholly external to myself, for

instance, or even one internal to myself, such as a nerve impulse, volition, or whatnot—then I cannot regard the behavior as being an act of mine, unless I further believed that I was the cause of that external or internal event."¹⁰

Chisholm tries to lessen the air of mystery surrounding immanent causation by appealing to eighteenth-century Scottish philosopher Thomas Reid, who is generally regarded as the father of modern agent-cause theories. Reid argued that the notion of agent-causation, far from being derivable from, or reducible to, causation in terms of events, is more fundamental than event-causation. Only by understanding our own causal efficacy as agents can we grasp the notion a *cause* at all: the notion of cause, he says, "may very plausibly be derived from the experience we have . . . of our own power to produce certain effects."¹¹ We then extend this power from ourselves to other things in the world. But our understanding of causal power comes first from our own experience as agents. So agent-causation may be difficult to understand, according to Reid. But we must believe in it nevertheless because we have direct experience of it in our daily lives; and the concept of event-causation is derived from that of agent-causation, not the other way around. As Chisholm says, taking his cue from Reid, "if we did not understand the concept of immanent causation, we would not understand that of transeunt causation."¹²

Reid and Chisholm may be right that we get our first ideas of causal power from our own experience of agency. Some psychological studies support this idea. But this fact alone does not eliminate the problems surrounding their agent-causal view. The first problem is this: how can we know from the immediate experience of our own agency alone that our actions are not determined by events (some of which may be hidden from us)? We may feel this is not so. We may feel, as Taylor says, that *we*, as agents, are the only determiners of our actions. But how can we be sure? For agent-causalists to say that choices or actions that are immanently caused by agents cannot *by their very nature* be caused by prior events seems to answer this problem by stipulation. In saying such a thing, agent-causalists would seem to be defining immanent causation so that it cannot in principle be caused by other events. If so, they would be getting the result they want for free rather than by honest toil.

5. Agent-causation, Regresses, and Randomness

But for the sake of argument, suppose we grant their stipulation: the immanent causing of an action or event cannot by its nature be determined or caused by other events. Then a second problem arises: if agent-causal

events are not determined or caused, are they random? Does the agent-causal theory really eliminate the problem of *randomness* or *arbitrariness* about undetermined free choices? Recall how that problem was posed: if Mike may have chosen to vacation in Hawaii or Colorado, given all the same prior mental and physical circumstances leading to his choice, including exactly the same prior thought processes, why wouldn't his choice of one or the other, Hawaii or Colorado, have been random or arbitrary? Agent-causalists respond that the choice would not have occurred merely randomly or arbitrarily, "out of the blue," so to speak (even though it was undetermined by prior circumstances) because Mike, the *agent*, would have immanently caused whichever choice was made in a way that could not be fully explained by, or reduced to, causation by prior circumstances.

But does this really solve the problem of randomness or arbitrariness? If it would have been irrational, inexplicable, random, or arbitrary for Mike to choose to vacation in Colorado, given the same mental circumstances and at the end of the same deliberation that led him to favor and choose Hawaii, why would it not have been equally irrational, random, arbitrary and so on, for Mike to *agent- (or immanently) cause* the choice to vacation in Colorado (in these same mental circumstances and at the end of the same deliberation that led him to favor and choose Hawaii)? The problem of randomness or arbitrariness, rather than being solved, seems to be merely transferred from the randomness and arbitrariness of the *choices* to the randomness and arbitrariness of *agents'- (immanently)-causing-the-choices*.

Similar questions arise when we consider problems about luck and chance. John succumbed to temptation and chose to arrive at his meeting late. In exactly the same circumstances, John* overcame temptation and chose to arrive on time. According to the Luck Objection, if there is nothing about John's and John*'s powers, capacities, states of mind, moral character, and the like leading up to their choices that explains why John chose one way and John* another, then the difference is just a matter of luck. John got lucky in his attempt to overcome temptation, while John* did not.

Agent-causalists respond that merely because the choices of John and John* were not caused by prior events does not mean they merely occurred out of the blue, uncaused by *anything*. The choices were caused, not by prior events, but by the agents. John agent-caused his choice to arrive late (in a direct or immanent way that could not be explained in terms of causation by prior events) and John* agent-caused his choice to arrive on time in a similarly direct manner. So it was up to them which choice occurred.

But is the Luck Objection really answered by this argument? If it is a matter of luck or chance that John* chose to overcome temptation and John did not, why is it not equally a matter of luck or chance that John* (immanently) *agent-caused* the choice to overcome temptation while John did not? Since the immanent agent-causing of one choice rather than another is also undetermined by prior circumstances, then there is nothing about John's and John*'s powers, capacities, states of mind, and other prior circumstances that explains why they immanently *agent-caused* different choices. It seems that problems about luck or chance, like problems about randomness and arbitrariness, are merely transferred from the choices to the *agent-causing-of-the-choices* without being solved.

Chisholm is aware of these difficulties. He argues that to be consistent with their general strategy, agent-causalists should respond that the *agent-causing-of-the-choices* is not caused by prior events, but neither does it occur by luck or chance. There is a third option: the agent-causing of the choices is itself immanently caused by the agent. Chisholm realizes that this response unfortunately seems to give rise to an infinite regress: if John (or John*) is the agent-cause of his choice, he is also the agent-cause of his being the agent-cause of his choice and also the agent-cause of his being the agent-cause of his being the agent-cause of his choice, and so on indefinitely. This is an unhappy consequence to say the least: it seems that an infinite series of agent-causings would be needed for each free choice. But Chisholm bites the bullet and accepts this consequence anyway because he thinks that if the regress stopped at any point, it would not be clear that the first immanent causing was "up to the agent" rather than occurring merely randomly or by chance. To make this infinite series of immanent causings seem less a violation of common sense, Chisholm adds that the agents need not be *aware* of all these agent-causings, for the doctrine of agent-causation does not require that agents be aware of all the events they agent-cause.

Nonetheless, most philosophers, and most agent-causalists themselves, are not comfortable with postulating an infinite series of agent-causings, as Chisholm does. Fortunately, there is another alternative open to them that most agent-causalists have preferred. "Chisholm's mistake," many of them say,

is assuming that agent-causation is an event like any other event that must either be caused or occur randomly. The agent-causal relation is unique and cannot be treated like any other event or occurrence. To ask the question 'if the agent-causal relation is not caused, why doesn't it occur merely randomly or by chance?' is to show you do not really understand what the agent-causal relation is. Immanent agent-causation is not the sort of thing that *can* in principle occur randomly or by

chance, any more than it can in principle be caused. For the agent-causal relation just *is* the agent's exercising conscious control over an event; and an agent's exercising conscious control over an event is not the sort of thing that happens out of the blue, by chance or accident. For by its nature it is up to the agent. We do not need a further agent-causing to explain it.

This response avoids Chisholm's regress, to be sure. But if agent-causalists respond in this way, it seems they are once again solving the problems about libertarian free will by stipulation. In response to the objection that for all we know immanent agent-causation might be determined by hidden causes, they insist that immanent agent-causation is not the sort of thing that could in principle be caused or determined by prior events or circumstances. Now, in response to the randomness and luck objections, they add that the agent-causal relation is not the sort of thing that could in principle occur randomly or by chance either, since it is the agent's consciously controlling something.

To many critics of libertarianism, this solution looks like solving the Libertarian Dilemma—either determinism or mere chance—by a *double* stipulation, by introducing a special agent-causal relation defined in such a way that it (1) cannot by its nature be determined, but (2) cannot by its nature be random either. One can see why many critics of libertarianism think that agent-cause theories either lead to infinite regresses or solve the problems about libertarian free will by defining them out of existence (for "free" rather than by honest toil). Gary Watson states this criticism in the following words:

All we know of this [agent-causal] relation is that it holds between an agent and an event when the agent is the responsible agent of that event, and the event is uncaused by other events. . . . Agent-causation meets [these] conditions . . . by stipulation. But the challenge is to say what this [agent-causal] relation amounts to in such a way as to give some reason for thinking it is empirically possible. 'Agent-causation' simply labels, not illuminates, what the libertarian needs.¹³

Watson's point is that if agent-causalists are to do more than merely label what libertarians need, they must say more about the nature of agent-causation and do more to show how such a thing is empirically possible. Failing to do that, agent-causalist solutions to the free will problem will remain as mysterious as Kantian and dualist solutions. In the next chapter, we will consider what other strategies are available to libertarians, agent-causalists, and others to make sense of the "deeper" freedom of the will they believe in.

Suggested Reading

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



Actions, Reasons, and Causes

1. Simple Indeterminism

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

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

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