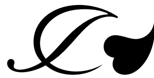
A Contemporary Introduction to Free Will



ROBERT KANE
University of Texas at Austin

New York ◆ Oxford
OXFORD UNIVERSITY PRESS
2005

Oxford University Press

Oxford New York
Auckland Bangkok Buenos Aires Cape Town Chennai
Dar es Salaam Delhi Hong Kong Istanbul Karachi Kolkata
Kuala Lumpur Madrid Melbourne Mexico City Mumbai Nairobi
São Paulo Shanghai Taipei Tokyo Toronto

Copyright © 2005 by Oxford University Press, Inc.

Published by Oxford University Press, Inc.
198 Madison Avenue, New York, New York 10016
www.oup.com

Oxford is a registered trademark of Oxford University Press

All rights reserved. No part of this publication may be reproduced,
stored in a retrieval system, or transmitted, in any form or by any means,
electronic, mechanical, photocopying, recording, or otherwise,
without the prior permission of Oxford University Press.

ISBN-13: 978-0-19-514969-2 (alk. paper)—
ISBN-13: 978-0-19-514970-8 (pbk.: alk. paper)
ISBN-10: 0-19-514969-6 (alk. paper)—ISBN-10: 0-19-514970-X (pbk.: alk. paper)

Printing number: 9 8 7 6 5 4 3 2 1

Printed in the United States of America
on acid-free paper

Contents

ACKNOWLEDGMENTS **vii**

1 The Free Will Problem **1**

2 Compatibilism **12**

3 Incompatibilism **23**

4 Libertarianism, Indeterminism, and Chance **32**

5 Minds, Selves, and Agent Causes **40**

6 Actions, Reasons, and Causes **53**

7 Is Free Will Possible? Hard Determinists and Other Skeptics **67**

8 Moral Responsibility and Alternative Possibilities **80**

9 Higher-order Desires, Real Selves, and New Compatibilists **93**

10 Reactive Attitude Theories **107**

11 Ultimate Responsibility **120**

12 Free Will and Modern Science **132**

13 Predestination, Divine Foreknowledge, and Free Will **147**

14 Conclusion: Five Freedoms **163**

NOTES **175**

INDEX **185**

CHAPTER 2

Compatibilism

1. Introduction

The view that there is really is no conflict between determinism and free will—that free will and determinism are compatible—is known as *compatibilism;* and it is the first view about free will we shall consider. Compatibilism has become an increasingly popular doctrine in modern philosophy because it provides what seems to be a neat, simple solution to the free will problem. If there really is no conflict between free will and determinism, as compatibilists say, then the age-old problem of free will is resolved in one fell swoop.

Compatibilism was held by some ancient philosophers, like the Stoics, and perhaps Aristotle too, according to some scholars. But it has become especially popular since the seventeenth century. Influential philosophers of the modern era, such as Thomas Hobbes, John Locke, David Hume, and John Stuart Mill, were compatibilists. They saw compatibilism as a way of reconciling ordinary experience of being free with scientific views about the universe and human beings. Compatibilism remains popular among philosophers and scientists today for similar reasons. If compatibilists are right, we can have both freedom and determinism, and need not worry that future science will somehow undermine our ordinary conviction that we are free and responsible agents.

This is a comforting thought. But is compatibilism believable? In my experience, most persons resist the idea that free will and determinism might be compatible when they first encounter it. The idea that determinism might be compatible with freedom and responsibility looks at first like a "quagmire of evasion," as William James called it, or a "wretched subterfuge" as Kant called the compatibilism of Hobbes and Hume. If

compatibilism is to be taken seriously by ordinary persons, they have to be talked out of this natural belief in the incompatibility of free will and determinism by means of philosophical arguments; and supplying such arguments is what compatibilists try to do.

2. Freedom as the Absence of Constraints

The first step in the compatibilists' argument is to ask us to reflect on what we ordinarily mean by saying actions or choices are "free." What does it mean to say I am free to take the bus this morning? It does not mean I will actually take the bus, for I may choose not to take it. But I am free to take the bus, if I have the *power* or *ability* to take it, should I want or decide to do so. Freedom then is, first of all, a power or ability to do something, a power I may or may not choose to exercise.

Second, this power or ability, which is my freedom, entails that there are no *constraints* or *impediments* preventing me from doing what I want to do. I would not be free to take the bus if various things prevented me: such as being in jail or if some one had tied me up (physical restraint); or if someone were holding me at gunpoint, commanding me not to move (coercion); or if I were paralyzed (lack of ability); or if buses were not running today (lack of opportunity); or if fear of crowded buses compelled me to avoid them (compulsion), and so on.

Putting these thoughts together, compatibilists argue that to be free, as we ordinarily understand it, is (1) to have the *power* or *ability* to do what we want or desire to do, which in turn entails (2) an *absence of constraints* or impediments (such as physical restraints, coercion, and compulsion) preventing us from doing what we want. Let us call a view that defines freedom in terms of 1 and 2 "classical compatibilism." Most traditional compatibilists, such as Hobbes, Hume, and Mill, were classical compatibilists in this sense. Hobbes stated the view succinctly, saying a man is free when he finds "no stop in doing what he has the will, desire or inclination to do."[1] And Hobbes noted that if this is what freedom means, then freedom is compatible with determinism. For, as he put it, there may be no constraints or impediments preventing persons from doing what they "will or desire to do," even if it should turn out that what they will or desire was determined by their past.

But doesn't freedom also require alternative paths into the future, and hence the freedom *to do otherwise?* How do classical compatibilists account for the freedom to do otherwise? They begin by defining the freedom to do otherwise in terms of the same conditions 1 and 2. You are free to do otherwise than take the bus if (1) you have the power or ability to

avoid taking it, which entails (2) that there are also no constraints preventing you from *not* taking the bus, if you wanted to (no one is holding a gun on you, for example, forcing you to get on the bus.)

Of course, an absence of constraints preventing you from doing otherwise does not mean you will actually do otherwise. But, for classical compatibilists, the freedom to do otherwise does mean that you *would* have done otherwise (nothing would have stopped you) *if* you had wanted or desired to do otherwise. And they argue that if the freedom to do otherwise has this *conditional* or *hypothetical* meaning (you *would* . . . , *if* you wanted to), then the freedom to do otherwise would also be compatible with determinism. For it may be that you *would* have done otherwise *if* you had wanted to, even though you did not in fact want to do otherwise, and even if what you wanted to do was determined.

3. Freedom of Will

Is this classical compatibilist account of freedom plausible? It does seem to capture the *surface freedoms* discussed in chapter 1. Surface freedoms, you may recall, were those everyday freedoms to buy what we want, walk where we please, take buses when we want to, without anything preventing us. These everyday freedoms do seem to amount to (1) the power or ability to do what we want (and the power to have done otherwise, *if* we had wanted to) and (2) doing so without any constraints or impediments getting in our way. But if the classical compatibilist analysis of freedom does capture these surface freedoms of *action* discussed in chapter 1, does it also capture the "deeper" freedom of the *will?*

Classical compatibilists respond to this question in two ways. First, they say:

> It all depends on what you mean by "freedom of will." In one sense, freedom of will has a perfectly ordinary meaning. For most of us, it means *freedom of choice* or *decision*. But freedom of choice or decision can be analyzed in the same way that we compatibilists analyze freedom of action generally. You are free to *choose* to lend money to a friend, for example, if (1) you have the power or ability to *choose* to lend the money in the sense that (2) no constraints would prevent you from making the choice, *if* you wanted to, and, in addition, nothing would have prevented you from *choosing otherwise* (choosing not to lend the money), if you had wanted to choose otherwise.

In short, compatibilists say that free choices or decisions can be treated like free actions of other kinds. For, choices or decisions can be subject to

constraints just like other kinds of actions; and when choices or decisions are subject to constraints, they are also not free. For example, you might have been brainwashed or hypnotized, so that you could not have chosen otherwise (chosen not to lend money), even *if* you wanted to. Conditions such as brainwashing and hypnosis are two further constraints that can take away freedom; and they sometimes take away even the freedom to *choose* what we would otherwise have wanted to choose. When brainwashing or hypnosis do this they take away our freedom of *will*.

Here is another example of constraint on choices or decisions. If a man holds a gun to your head and says "Your money or your life," he is giving you a choice of sorts. You can choose to hand over your money or take a chance on losing your life. But in another sense, the man has not given you any *real* choice at all, if you believe he is serious. For the prospect of losing your life is so horrible this is no choice at all. Your choice to hand over the money is therefore not really free. It is *coerced;* and coercion is a constraint on your freedom of choice or freedom of will. The thief's actions have kept you from making the choice you really wanted to make, which was to keep both money *and* life.

So the first response of compatibilists regarding "freedom of will" is to say that if freedom of will means what we usually mean by it—*unconstrained freedom of choice or decision*—then freedom of will can also be given a compatibilist analysis. You have freedom of will when nothing would have prevented you from choosing *or* from choosing otherwise *if* you had wanted to; and if this is what freedom of will means, they argue, then freedom of will (as well as freedom of action) is consistent with determinism.

4. If the Past Had Been Different

But compatibilists are aware that many persons are not going to be satisfied with this account of free will as mere unconstrained choice or decision. So they have a second response.

> If you are still not satisfied with the above account of freedom of will, then it is no doubt because you are thinking of free will in some further sense than simply the ability to choose or decide *as* you will without constraint. You must be thinking of freedom of will in something like the 'deeper' sense of free will of chapter 1—as a kind of *ultimate* control over what you will or want in the first place: A control incompatible with your will's being determined by any events in the past over which you did not have control. Now we compatibilists obviously can't

capture *that* deeper sense of freedom of will, no matter what we do, because it is incompatible with determinism. But, as compatibilists, we believe that any so-called deeper freedom of the will—or any kind of free will that requires indeterminism—is incoherent anyway. No one *could* have a freedom of will of such a deeper kind.

Why do compatibilists believe that any kind of deeper freedom of will that requires indeterminism must be incoherent? Well, if determinism means (as it does): *same past, same future,* then, the denial of determinism—indeterminism—must mean: *same past, different possible futures*. (Think of the garden of forking paths of chapter 1.) But if that is what indeterminism means—same past, different possible futures—indeterminism has some odd consequences regarding free choices. Consider Molly again deliberating about whether to join the law firm in Dallas or the one in Austin. After much thought, let us say, Molly decided that the Dallas firm was a better one for her career plans and she chose it. Now if her choice was undetermined, she might have chosen differently (she might have chosen the Austin firm instead), *given the same past*—since that is what indeterminism requires: same past, different possible futures. But note what this requirement means in Molly's case: exactly the same prior deliberation, the same thought processes, the same beliefs, desires, and other motives (not a sliver of difference!) that led to Molly favoring and choosing the Dallas firm *might have issued in her choosing the Austin firm instead*.

That senario makes no sense, say compatibilists. It would be senseless and irrational for Molly to choose the Austin firm, given exactly the same motives and prior process of reasoning that *in fact* led her to believe the Dallas firm was the better one for her career. To say that Molly "could have chosen otherwise"in these circumstances must mean something else, say compatibilists—something like the following: *if* Molly had had *different* beliefs or desires, or had reasoned differently, or *if* other thoughts had entered her mind before she chose the Dallas firm, *then* she might have come to favor the Austin firm instead and chosen it. But this more sensible interpretation of "could have done otherwise," say compatibilists, means only that Molly would have done otherwise, if things had been different—if *the past had been different in some way.* And such a claim, they insist, does not conflict with determinism. In fact, this interpretation of "could have chosen otherwise" perfectly fits the classical compatibilists' *conditional* or *hypothetical* analysis—"Molly could have chosen otherwise" means "She *would* have chosen otherwise, *if* she had wanted to (if her mind-set had been different in some way). And such a hypothetical interpretation of "could have chosen otherwise" is, as we have seen, compatible with determinism.

One's first thought when encountering this argument is that there must be some way around the conclusion that if Molly's choice is undetermined, she must have been able to choose otherwise "given exactly the same past." But in fact there is no easy way around this conclusion. For indeterminism, which is the denial of determinism, *does* mean "different possible futures, given the same past." In the diagram of forking paths of chapter 1, the single line going back into the past is just that: a single line indicating "same past"; while the multiple lines going into the future represent "different possible futures." By contrast, determinism means only one line into the future. If Molly really is free to choose different options at any time during her deliberation, and her choice is not determined, then she must be able to choose *either* path (the Dallas firm or the Austin firm), given the *same* past up to the moment when she chooses.

You can't cheat here by suggesting that if the past had been a *tiny bit* different, then Molly might have chosen differently (chosen the Austin firm). *Determinists* and *compatibilists* can say this: for they insist that Molly might have sensibly and rationally chosen otherwise only if the past had been different in some way (however small the difference). But persons who believe free choices cannot be determined must say that Molly may have chosen different possible futures, given the same past at the time she did choose. And this does seem to make choosing otherwise in the same circumstances arbitrary and irrational.

To sum up: compatibilists have a twofold response to the objection that their view accounts only for freedom of action but not for freedom of will. On the one hand, they say, if "freedom of will" means what we ordinarily mean by free *choices* or *decisions* (those that are uncoerced and unconstrained), then freedom of will can also be given a compatibilist analysis and can thus be seen to be compatible with determinism. On the other hand, if "freedom of will" has a stronger meaning—if it refers to some kind of "deeper" freedom of the will that is not compatible with determinism—then that deeper freedom of will is incoherent and is not something we can have anyway.

5. Constraint, Control, Fatalism, and Mechanism

So far, the compatibilist argument has been that people believe determinism conflicts with free will because they have confused ideas about *freedom*. But compatibilist arguments about freedom of action and will are only half of the compatibilists' case. They also argue that people mistakenly believe determinism and free will conflict because they also have confused ideas about *determinism*. Determinism, compatibilists insist, is not

the frightful thing we think it is. People believe determinism is a threat to freedom because they commonly confuse determinism with a host of other things that are a threat to freedom. But determinism does not imply these other threatening things, according to compatibilists. For example, they say:

1. "Don't confuse *determinism* with *constraint, coercion,* or *compulsion*." Freedom *is* the opposite of constraint, coercion, and compulsion compatibilists insist; but it is not the opposite of determinism. Constraint, coercion, and compulsion act *against* our wills, preventing us from doing or choosing what we want. By contrast, determinism does *not* necessarily act against our wills; nor does it always prevent us from doing what we want. Causal determinism, to be sure, *does* mean that all events follow from earlier events in accordance with invariable laws of nature. But, say compatibilists, it is a mistake to think that laws of nature *constrain* us. According to A. J. Ayer (a noted twentieth-century compatibilist), many people think freedom is inconsistent with determinism because they have a mistaken image of natural causes or laws of nature "overmastering" us, forcing us against our wills. But, in fact, the existence of laws of nature indicates only that certain events follow others according to regular patterns. To be governed by laws of nature is not to be in chains.

2. "Don't confuse *causation* with *constraint*." Compatibilists also insist that it is constraints, not mere *causes* of any kind, that undermine freedom. Constraints *are* causes, but they are causes of special kinds: impediments or hindrances to our doing what we want, such as being tied up or paralyzed. Not all causes are impediments to freedom in this sense. In fact, some causes, such as muscular strength or inner strength of will, actually *enable* us to do what we want. It is therefore a mistake to think that actions are unfree simply because they are caused. Whether actions are free or not depends on what *kinds* of causes they have: some causes enhance our freedom, while other causes (i.e., constraints) hinder our freedom.

It is a further mistake, say compatibilists, to think that, when we act or choose freely in accordance with our wills, our actions are entirely *uncaused*. To the contrary, our free actions are caused by our characters and motives; and this state of affairs is a good thing. For if actions were not caused by our characters and motives, we could not be held responsible for the actions. They would not be *our* actions. This point was made in a well-known passage by perhaps the most influential classical compatibilist, David Hume:

> Where [actions] proceed not from some *cause* in the character and disposition of the person who performed them, they can neither redound to his honour, if good; nor infamy, if evil. . . . The person is not answerable for them;

> and as they proceeded from nothing in him that is durable and constant . . . it is impossible he can, upon their account, become the object of punishment or vengeance.[2]

Classical compatibilists follow Hume in saying that responsible actions cannot be uncaused; such actions must have the right kinds of causes—causes that come from inside our selves and express our characters and motives, rather than causes imposed upon us against our wills. It is a mistake to think that free will and determinism are not compatible because free actions should be uncaused. Free actions are *unconstrained,* not *uncaused*.

3. "Don't confuse *determinism* with *control* by other agents." Compatibilists can concede (and often do concede) that it *does* count against our freedom if we are controlled or manipulated by other *persons*. That is why sci-fi utopias, like *Brave New World* and *Walden Two,* where people are controlled by behavior engineers or neurochemists, seem to undermine human freedom. But compatibilists insist that determinism by itself does not necessarily imply that any other persons or agents are controlling our behavior or manipulating us.

Nature by itself "does not control us," says compatibilist Daniel Dennett, since nature is not an agent.[3] What is objectionable about control by other agents, Dennett argues—whether they be behavioral engineers or con men—is that other persons are using us as means to their ends, lording it over us and making us conform to their wishes. We resent this kind of interference. But merely being determined does not imply that any other *agents* are interfering with us or using us in this way. So compatibilists can reject Brave New World and Walden Two scenarios, says Dennett, without giving up their belief that determinism is consistent with freedom and responsibility.

4. "Don't confuse *determinism* with *fatalism*." This is one of the most common confusions in free will debates. Fatalism is the view that whatever is going to happen, is going to happen, *no matter what we do*. Determinism alone does not imply such a consequence. What we decide and what we do would make a difference in how things turn out—often an enormous difference—even if determinism should be true. This important point was made by another influential classical compatibilist, John Stuart Mill:

> A fatalist believes . . . not only that whatever is about to happen will be the infallible result of causes that precede it [which is what determinists believe], but moreover that there is no use in struggling against it; that it will happen however we may strive to prevent it. . . . [Thus, fatalists believe that a man's]

> character is formed *for* him, and not *by* him; therefore his wishing it was formed differently is of no use; he has no power to alter it. This is a grand error. He has, to a certain extent, a power to alter his character. Its not being, in the ultimate resort, formed for him, is not inconsistent with its being, in part, formed *by* him as one of the immediate agents. His character is formed by his circumstances . . . but his own desire to mold it in a particular way is one of those circumstances, and by no means the least influential.[4]

Determinism, Mill is saying, does not imply that we have no influence on how things turn out, including the molding of our characters. We obviously do have such an influence, and determinism alone does not rule it out. Believing in fatalism, by contrast, can have fatal consequences. A sick man may excuse himself for not seeing a doctor saying: "If your time is up, it doesn't matter what you do about it." Or a soldier may use a familiar line for not taking precautions: "There's a bullet out there with your name on it. When it comes, you will not be able to avoid it, no matter what you do." Mill is saying that such fatalist claims do not follow merely from determinism. To think they do is a "grand error."

The claims of the sick man and the soldier are in fact examples of what the ancient philosophers called the "lazy sophism" ("sophism" meaning a fallacy of reasoning). The proper answers to the sick man and the soldier would be, "*Whether* your time is now up may depend in great part on whether you see a doctor; and *whether* any bullet out there right now has your name on it may depend on what precautions you take. So instead of sitting around doing nothing, see a doctor and take precautions." This is the response that compatibilists, such as Mill, would give to the "lazy sophism." Believing that determinism is compatible with freedom, they would say, should not make you a fatalist. Indeed this belief should convince you that your life is to some extent in your own hands, since how you deliberate can still make a difference in your future, even if determinism should turn out to be true.

Sometimes our deliberations do not matter to our fate, but not always. For example, Dennett describes a despairing man who jumps off a bridge intending to commit suicide. Halfway down, the man deliberates again, and thinks of life from a different perspective, deciding that perhaps suicide isn't a good idea after all. Now *this* man's deliberation no longer does matter to his fate. But ordinarily when we deliberate we are not in such desperate straits. Indeed, conditions like this man's are rare. Most of the time, say compatibilists, our deliberations do affect our future, even if determinism should be true.

5. "Don't confuse *determinism* with *mechanism*." Another common confusion, according to compatibilists, is to think that if determinism were true, we would all be machines, running mechanically, like watches,

robots, or computers. Or, alternatively, we would be like amoebae or insects and other lower creatures responding automatically, and with a fixed set of responses, to the stimuli of our environment. But, compatibilists insist, none of these consequences follows from determinism either.

Suppose it should turn out that the world is determined. There would still be an enormous difference between human beings, on the one hand, and amoebae and insects, or machines and robots, on the other. Unlike machines (even complex machines like computers) or robots, we humans have an inner conscious life of moods and feelings, and we react to the world accordingly. And unlike amoebae, insects, and other such creatures, we do not just react to the environment instinctually and in automatic ways. We reason and deliberate, question our motives, reflect on our values, make plans about the future, reform our characters, and make promises to others that we then feel obligated to keep.

Determinism does not rule out any of these capacities, say compatibilists, and they are the capacities that make us free and responsible beings, capable of moral action—as machines and insects are not. Determinism does not necessarily imply mechanical, inflexible, or automatic behavior either. Determinism is consistent with a whole spectrum of complexity and flexibility of behavior in living things, from the simplest amoeba all the way to human beings. The complexity and degrees of freedom of creatures in the world, from amoebae to humans, might differ incredibly, yet all these properties might be determined.

6. Assessing Classical Compatibilism

In summary, classical compatibilists say that our natural belief in the incompatibility of free will and determinism rests on confusions of two kinds—confusions about the nature of *freedom* and confusions about the nature of *determinism*. Once these confusions have been cleared up, they insist, we should see there is no necessary conflict between freedom and determinism. To assess the classical compatibilists' position, one must therefore ask whether their account of freedom really does capture what we mean by freedom of will and action; and one must ask whether the belief that determinism conflicts with free will does rest on confusions about determinism. Both these questions will be considered in the next chapter.

It is worth noting in conclusion, however, that classical compatibilists do seem to be right about certain things, whatever the final judgment may be about their view. They would appear to be right, for example, in saying determinism *in and of itself* does not imply *constraint, control by other agents, fatalism,* or *mechanism*. These *would* indeed rule out free will, but determinism does not necessarily imply them, and it would be a mistake to

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).

CHAPTER 3

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.

CHAPTER 4

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

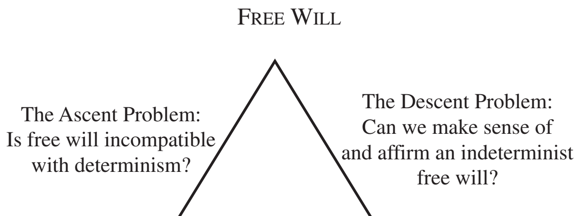


Figure 4.1 Incompatibilist Mountain and the Libertarian Dilemma

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 Bogeyman

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 that 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).

CHAPTER 5

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.

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 a 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).

CHAPTER 7

Is Free Will Possible? Hard Determinists and Other Skeptics

1. *Oklahoma City and Columbine*

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

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

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

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

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

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

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

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

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

2. Hard Determinism

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

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

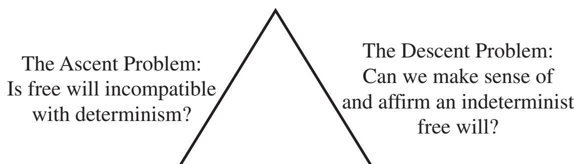


Figure 7.1 Incompatibilist Mountain and the Libertarian Dilemma

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

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

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

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

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

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

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

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

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

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

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

Strawson then approvingly quotes Friedrich Nietzsche, who said:

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

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

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

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

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

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

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

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

4. Living Without Free Will: Crime and Punishment

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

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

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

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

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

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

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

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

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

> for the disease. If a child is a carrier of the Ebola virus by virtue of its being passed on to her at birth from her parent, quarantine is nevertheless intuitively legitimate.
>
> Furthermore, if we have the right to "quarantine" criminals, we have the right to tell people in advance that they will be isolated from society if they commit crimes. . . . This publicity itself has a powerful deterrent effect.[6]

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

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

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

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

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

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

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

6. Illusion and Free Will

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

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

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

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

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

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

Suggested Reading

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

Notes

Chapter 1

1. For discussion of various interpretations of quantum physics in relation to free will, see essays by Robert Bishop and David Hodgson in Robert Kane, ed., *The Oxford Handbook of Free Will* (Oxford: Oxford University Press, 2002).
2. For example, Ted Honderich, *How Free Are You?* (Oxford: Oxford University Press, Clarendon Press, 1993).

Chapter 2

1. Thomas Hobbes, *Leviathan* (Indianapolis: Bobbs-Merrill, 1958), p. 108.
2. David Hume, *A Treatise on Human Nature* (Oxford: Oxford University Press, Clarendon Press, 1960), p. 411. An excellent account of Hume's compatibilist view is Paul Russell's *Freedom and Moral Sentiment* (New York: Oxford University Press, 1995).
3. Daniel Dennett, *Elbow Room: The Varieties of Free Will Worth Wanting* (Cambridge, MA: MIT Press, 1984), p. 61.
4. John Stuart Mill, *A System of Logic* (New York: Harper & Row, 1874), p. 254.

Chapter 3

1. Peter van Inwagen, *An Essay on Free Will* (Oxford: Oxford University Press, Clarendon Press, 1983), p. 16. The fuller version of the argument presented in the remainder of this section is my own interpretation of van Inwagen's argument.
2. van Inwagen, 1983; Carl Ginet, *On Action* (Cambridge: Cambridge University Press, 1990). I have put their response in my own words.
3. Michael McKenna, "Compatibilism," in Edward N. Zalta, ed., *The Stanford Encyclopedia of Philosophy,* online edition: http://plato.Stanford.edu/archives/sum2004/entries/compatibilism/. An objection of this kind was originally made by Keith Lehrer.

Chapter 4

1. This designation is Gary Watson's, in the second edition of *Free Will* (Oxford: Oxford University Press, 1982), p. 10.
2. Arthur Schopenhauer, *Prize Essay on the Freedom of the Will,* ed. with an introductions by Gunter Zoller. Translated by E.J.F. Payne. (Cambridge: Cambridge University Press, 1960), p. 47.
3. G. W. F. Leibniz, *Selections* (New York: Scribner's, 1951), p. 435.
4. This objection has been made by Galen Strawson, Alfred Mele, Bernard Berofsky, Bruce Waller, Richard Double, Mark Bernstein, and Ishtiyaque Haji. Statements of it can be found in the suggested readings for chapter 1 and for this chapter.
5. Bruce Waller, "Free Will Gone Out of Control: A Critical Study of R. Kane's *Free Will and Values*," *Behaviorism* 16 (1988): 149–67; quotation, p. 151.
6. Alfred Mele, "Review of Kane, *The Significance of Free Will,*" *Journal of Philosophy* 95 (1998): 581–4; quotation, pp. 582–83.

Chapter 5

1. Simon Blackburn, *Think* (Oxford: Oxford University Press, 1999), p. 89.
2. Immanuel Kant, *Critique of Pure Reason* (London: Macmillan, 1958), pp. 409–15.
3. Kant, *Foundations of the Metaphysics of Morals* (Indianapolis: Bobbs-Merrill, 1959), pp. 64–72.
4. R. M. Chisholm, "Human Freedom and the Self," in Gary Watson, ed., 2nd ed., *Free Will* (Oxford: Oxford University Press, 2003),

pp. 24–35. Also in Robert Kane, ed., *Free Will,* (Oxford: Blackwell Publishers, 2002), pp. 47–58, and in Laura Waddell Ekstrom, ed., *Agency and Responsibility: Essays on the Metaphysics of Freedom* (Boulder, Co: Westview Press, 2001), pp. 126–37.

5. Ibid., p. 30.
6. Ibid., p. 31.
7. Richard Taylor, *Metaphysics* (Englewood Cliffs: Prentice-Hall, 1974), p. 56.
8. Chisholm, op. cit., p. 34.
9. Taylor, op. cit., p. 57
10. Taylor, op. cit., p. 55.
11. *The Works of Thomas Reid,* ed. W. Hamilton (Hildeshein: George Ulm, 1983), p. 599.
12. Chisholm, op. cit., p. 31.
13. Gary Watson, ed. *Free Will* (Oxford: Oxford University Press, 1982), p. 10.

Chapter 6

1. Carl Ginet, *On Action* (Cambridge: Cambridge University Press, 1990), 11ff.
2. Alfred Mele, *Motivation and Agency* (Oxford: Oxford University Press, 2003), pp. 42–43.
3. R. E. Hobart, "Free Will as Involving Determinism and Inconceivable Without It," *Mind* 32 (1934): 1–27; quotation, p. 5.
4. Timothy O'Connor, ed., *Persons and Causes: The Metaphysics of Free Will* (New York: Oxford University Press, 2000). pp. 85–95.
5. Stewart C. Goetz, "Review of O'Connor, *Persons and Causes,*" *Faith and Philosophy* 19 (2002): 116–20; p. 118. Also see his "A Non-causal Theory of Agency," *Philosophy and Phenomenological Research* 49 (1988): 303–16.
6. O'Connor, op. cit., pp. 85–95.
7. O'Connor, in Gary Watson ed., 2nd ed., *Free Will* (Oxford: Oxford University Press, 2003), pp. 271–72.
8. Ibid., p. 271.
9. Causal theorists of action include Donald Davidson, *Essays on Actions and Events* (Oxford: Oxford University Press, 1980), Alfred Mele, *Autonomous Agents: From Self-Control to Autonomy* (New York: Oxford University Press, 1995), and many others.
10. Randolph Clarke, *Libertarian Accounts of Free Will* (Oxford: Oxford University Press, 2003).
11. Gary Watson, ed., *Free Will* (Oxford: Oxford University Press, 1982).

12. Ginet, in Robert Kane, ed., *The Oxford Handbook of Free Will* (Oxford: Oxford University Press, 2002), p. 398.
13. O'Connor, op. cit., p. 79.
14. O'Connor, 2000; William Hasker, *The Emergent Self,* (Ithaca, NY: Cornell University Press, 1999).
15. Jan Cover and John O'Leary-Hawthorne, "Free Agency and Materialism," in D. Howard-Snyder and J. Jordan, eds., *Faith, Freedom and Rationality* (Lanham, MD: Rowman & Littlefield, 1996) make interesting arguments suggesting that this might be so.
16. See the suggested readings at end of chapter.

Chapter 7

1. Derk Pereboom, *Living Without Free Will* (Cambridge: Cambridge University Press, 2001), chapter 1.
2. Galen Strawson, *Freedom and Belief* (Oxford: Oxford University Press, 1986).
3. G. Strawson, "The Bounds of Freedom," in Robert Kane, ed., *The Oxford Handbook of Free Will* (Oxford: Oxford University Press, 2002), pp. 441–60; quotation, p. 444.
4. F. W. Nietzsche, in ibid.
5. Ted Honderich, *How Free Are You?* (Oxford: Oxford University Press, 1993).
6. Pereboom, 2001, p. 174.
7. Pereboom, "Living Without Free Will: The Case for Hard Compatibilism," in Kane, ed., *The Oxford Handbook of Free Will*, pp. 477–88; quotation, p. 486.
8. Ibid., p. 487.
9. Ibid., p. 485.
10. Saul Smilansky, *Free Will and Illusion* (Oxford: Oxford University Press, Clarendon Press, 2000).
11. Smilansky, "Free Will, Fundamental Dualism and the Centrality of Illusion," in Kane, ed., *The Oxford Handbook of Free Will,* pp. 489–505; quotation, pp. 498–89.
12. Ibid., p. 482.