

# Four Views on Free Will

# Great Debates in Philosophy

*Series Editor: Ernest Sosa*

Dialogue has always been a powerful means of philosophical exploration and exposition. By presenting important current issues in philosophy in the form of a debate, this series attempts to capture the flavor of philosophical argument and to convey the excitement generated by the exchange of ideas. Each author contributes a major, original essay. When these essays have been completed, the authors are each given the opportunity to respond to the opposing view.

## Personal Identity

*Sydney Shoemaker and Richard Swinburne*

## Consciousness and Causality

*D. M. Armstrong and Norman Malcolm*

## Critical Theory

*David Couzens Hoy and Thomas McCarthy*

## Moral Relativism and Moral Objectivity

*Gilbert Harman and Judith Jarvis Thomson*

## Atheism and Theism, Second Edition

*J. J. C. Smart and J. J. Haldane*

## Three Methods of Ethics

*Marcia W. Baron, Philip Pettit, and Michael Slote*

## Epistemic Justification

*Laurence Bonjour and Ernest Sosa*

## Four Views on Free Will

*John Martin Fischer, Robert Kane, Derk Pereboom, and Manuel Vargas*

# Four Views on Free Will

John Martin Fischer,  
Robert Kane,  
Derk Pereboom,  
and  
Manuel Vargas



© 2007 by John Martin Fischer, Robert Kane, Derk Pereboom, and Manuel Vargas

BLACKWELL PUBLISHING

350 Main Street, Malden, MA 02148–5020, USA

9600 Garsington Road, Oxford OX4 2DQ, UK

550 Swanston Street, Carlton, Victoria 3053, Australia

The right of John Martin Fischer, Robert Kane, Derk Pereboom, and Manuel Vargas to be identified as the Authors of this Work has been asserted in accordance with the UK Copyright, Designs, and Patents Act 1988.

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, except as permitted by the UK Copyright, Designs, and Patents Act 1988, without the prior permission of the publisher.

First published 2007 by Blackwell Publishing Ltd

1 2007

*Library of Congress Cataloging-in-Publication Data*

Four views on free will / John Martin Fischer . . . [et al.].

p. cm. — (Great debates in philosophy)

Includes bibliographical references (p. ) and index.

ISBN-13: 978-1-4051-3485-9 (hardback)

ISBN-13: 978-1-4051-3486-6 (pbk.)

1. Free will and determinism. I. Fischer, John Martin, 1952–

BJ1461.F68 2007

123'.5—dc22

2006026270

A catalogue record for this title is available from the British Library.

Set in 10 on 12.5 pt Caslon

by SNP Best-set Typesetter Ltd, Hong Kong

Printed and bound in Singapore

by Markono Print Media Pte Ltd

The publisher's policy is to use permanent paper from mills that operate a sustainable forestry policy, and which has been manufactured from pulp processed using acid-free and elementary chlorine-free practices. Furthermore, the publisher ensures that the text paper and cover board used have met acceptable environmental accreditation standards.

For further information on  
Blackwell Publishing, visit our website:  
[www.blackwellpublishing.com](http://www.blackwellpublishing.com)

# Contents

---

<i>Notes on Contributors</i>	vi
<i>Acknowledgments</i>	viii
A Brief Introduction to Some Terms and Concepts	1
1 Libertarianism <i>Robert Kane</i>	5
2 Compatibilism <i>John Martin Fischer</i>	44
3 Hard Incompatibilism <i>Derk Pereboom</i>	85
4 Revisionism <i>Manuel Vargas</i>	126
5 Response to Fischer, Pereboom, and Vargas <i>Robert Kane</i>	166
6 Response to Kane, Pereboom, and Vargas <i>John Martin Fischer</i>	184
7 Response to Kane, Fischer, and Vargas <i>Derk Pereboom</i>	191
8 Response to Kane, Fischer, and Pereboom <i>Manuel Vargas</i>	204
<i>Bibliography</i>	220
<i>Index</i>	224

# Notes on Contributors

---

---

**John Martin Fischer** is Distinguished Professor in the Department of Philosophy at the University of California, Riverside, where he holds a UC President's Chair. He is the author of *The Metaphysics of Free Will: An Essay on Control* (Blackwell, 1994); and, with Mark Ravizza, S.J., *Responsibility and Control: A Theory of Moral Responsibility* (1998). His collection of essays, *My Way: Essays on Moral Responsibility*, was published in 2006.

**Robert Kane** is University Distinguished Teaching Professor of Philosophy at the University of Texas at Austin. He is the author of seven books and over sixty articles on the philosophy of mind and action, ethics, the theory of value and philosophy of religion, including *Free Will and Values* (1985), *Through the Moral Maze: Searching for Absolute Values in a Pluralistic World* (1994), *The Significance of Free Will* (1996), *A Contemporary Introduction to Free Will* (2005), and a lecture series on audio and video tape entitled *The Quest for Meaning: Values, Ethics, and the Modern Experience*. He is editor of *The Oxford Handbook of Free Will* (2002) and of *Free Will* (Blackwell, 2002). The recipient of fifteen major teaching awards at the University of Texas, he was named an inaugural member of the University's Academy of Distinguished Teachers in 1995.

**Derk Pereboom** is Professor of Philosophy at the University of Vermont. He will join the Sage School of Philosophy at Cornell University in 2007. His book *Living Without Free Will* appeared in 2001, and he has published articles on free will, philosophy of mind, history of modern philosophy, and philosophy of religion.

**Manuel Vargas** is Associate Professor of Philosophy at the University of San Francisco. The author of various articles in ethics, philosophy of action, and Latin American philosophy, he has been awarded the American Philosophical Association's Prize in Latin American Thought (2004), and the N.E.H. Chair in the Humanities at the University of San Francisco (2005–2006).

# *Acknowledgments*

---

---

JMF: I'd like to thank Neal A. Tognazzini and John T. Maier for their helpful comments on my essay.

DP: Thanks from me to Manuel Vargas, Robert Kane, John Fischer, Seth Shabo, David Christensen, and Sarah Adler.

MV: Thanks to the other authors, and also to Eddy Nahmias, Shaun Nichols, and Dan Speak for very helpful feedback on my contributions to this book. Thanks to Neal Tognazzini for the index. Thanks also to Stephanie for lots of things.

RK: Thanks to the other three authors for their cogent comments, and to Manuel for his diligent organizational efforts.



# *A Brief Introduction to Some Terms and Concepts*

---

## **Basic Terms: Free Will, Moral Responsibility, and Determinism**

Perhaps the three most important concepts in philosophical work on free will are *free will*, *moral responsibility*, and *determinism*.

The notion of freedom at stake in philosophical discussions is usually distinguished from a variety of other freedom concepts, including things like religious and political freedom. Usually, **free will** is also treated as distinct from several other concepts associated with human agency, such as autonomy and authenticity. As we will see in the chapters that follow, there are many different ways of thinking about the nature of free will, and there are serious disagreements about what would constitute an adequate theory of free will. Much of the tradition has taken “free will” to be a kind of power or ability to make decisions of the sort for which one can be morally responsible, but philosophers have also sometimes thought that free will might be required for a range of other things, including moral value, originality, and self-governance. Two other claims often made about free will are hotly disputed among philosophers; and authors of this volume will take different sides on these claims. One is the claim that free will requires “alternative possibilities” or the power to do otherwise, and the other is the claim that free will requires that we are the “ultimate sources” of our free actions or the ultimate sources of our wills to perform free actions.

Important to many discussions of free will is the idea of **moral responsibility**. In the context of discussions of free will, moral responsibility is

often understood as a kind of status connected to judgments and/or practices of moral praise and blame. This meaning is distinct from another, perhaps more commonly used sense of responsibility: responsibilities as obligations (for example, when we talk about what responsibilities a parent has to a child). There are important connections between responsibility of the sort concerned with praise and blame and responsibility of the sort connected with obligations. However, philosophers writing on free will and moral responsibility are typically concerned with the former and not the latter.

**Determinism** is a third concept that is often important for philosophical discussions of free will. For present purposes, we can treat determinism as the thesis that at any time (at least right up to the very end) the universe has exactly one physically possible future. Something is deterministic if it has only one physically possible outcome.

It is important to bear in mind that a definition of determinism is just that – a characterization of what things would have to be like *if* things were deterministic. It does not follow that the universe is actually deterministic. Compare: “A creature is a gryphon if it has the hindquarters of a lion and the head and claws of an eagle.” Nothing about the definition of gryphon shows that there are such creatures in our universe. It simply tells us something about what sorts of things would count as gryphons. Similarly, to offer a definition of determinism does not show that the universe is deterministic. It only defines a term, and we may find that the term never properly applies to the world we live in.

When discussing these issues it is natural to wonder whether the world is deterministic. Most physicists and philosophers think that the answer is no, but the technical issues are extremely complex. Nevertheless, if we accept that the universe isn’t deterministic there are still good reasons to think about the compatibility of free will and determinism. First, it could turn out that future physicists conclude that the universe is deterministic, contrary to the contemporary consensus about at least quantum mechanics. It is notoriously difficult to predict how future science will turn out, and it might be useful to have an answer to the question in advance of the scientific issues getting sorted out. Second, even if the universe were not fully deterministic, determinism might hold locally (either as a matter of how local spacetime is constructed, or as a matter of how the physics for non-quantum physical objects operates). Third, we could be interested in whether free will is compatible with a broadly scientific picture of the universe. Since some aspects of the universe seem deterministic and others do not, we might ask if free will is compatible with determinism as a first step to answering the more general question of whether free will is compatible with a broadly scientific picture of the universe.

## Philosophical Options on the Free Will Problem

One particularly important issue for contemporary philosophers thinking about free will is whether we could have free will in a deterministic universe. Call this issue – whether free will could exist if the universe were deterministic – **the compatibility issue**. There is a long-standing tradition of dividing up the conceptual terrain in light of the main answers to the compatibility issue. Traditionally, **incompatibilists** are those who think that free will is incompatible with the world being deterministic. **Compatibilists**, conveniently enough, are those hold that free will is compatible with the universe being deterministic.

It is important to recognize that the compatibility issue is distinct from the issue of whether we have free will. You could be an incompatibilist, and maintain that we have do have free will. Or you might be an incompatibilist and think that we lack free will. (You could even think that irrespective of how the compatibility issue is settled, there are threats to free will apart from determinism.)

In the philosophical literature, **libertarianism** is the view that we have free will and that free will is incompatible with determinism. “Libertarianism” as it is used in the context of free will is distinct from libertarianism in political philosophy. (Indeed, “libertarianism” in the free will sense is the original meaning – it was only later appropriated as the label for a view in political philosophy.) One might be a libertarian in both political and free will senses, but you can be a libertarian about free will without being a libertarian in political philosophy. And, perhaps, you could also be a political libertarian without being a free will libertarian (although many political libertarians seem to also be free will libertarians).

Following Derk Pereboom, we will label as “**hard incompatibilism**” any view that holds that (1) incompatibilism is true and (2) we lack free will. Historically, most hard incompatibilists were what William James called **hard determinists**. (Indeed, Pereboom’s coining of the term “hard incompatibilism” reflects James’ older and narrower terminology.) Hard determinists think we lack free will *because the world is deterministic*. Contemporary hard determinists are few and far between. What is more common are views that hold that we have no free will irrespective of whether or not the world is deterministic, and views that hold that although freedom might be not be conceptually incompatible with determinism (or indeterminism, for that matter), we simply do not have it.

To summarize, then: A traditional way of dividing up the terrain concerns answers to the compatibility issue. The two main approaches are incompatibilism and compatibilism. We have been considering the incompatibilist fork, where the two main species of incompatibilism are libertarianism and hard

incompatibilism. Both forms of incompatibilism have further species we have not discussed in this brief introduction.

The remaining fork of the compatibility debate is **compatibilism**. There are many varieties of compatibilism. Some compatibilists have emphasized a particular understanding of “can,” others have emphasized a kind of identification with one’s motives or values, and others emphasizing the role of responsiveness to reasons. One influential variation, however, is the view that holds that responsibility is compatible with determinism, combined with agnosticism about whether free will understood in some particular way might not be compatible with determinism. This view is *semicompatibilism*, and its most prominent defender is John Martin Fischer.

Lastly, there are views that do not neatly fit the traditional taxonomy of incompatibilism and compatibilism. One such class of views is **revisionism**. The core idea of revisionism is that the picture of free will and moral responsibility embedded in commonsense is in need of revision, but not abandonment. That is, the revisionist holds that the correct account of free will and moral responsibility will depart from commonsense. As is the case with libertarianism, hard incompatibilism, and compatibilism, this view can take a variety of more specific forms.

For a different way to think about the relationship between the various views, see the grid below.

	Is commonsense thinking about free will and moral responsibility basically correct?	Is free will compatible with determinism?	Is moral responsibility compatible with determinism?	Do we have free will?
Libertarianism	Yes	No	No	Yes
Compatibilism	Yes	Yes (although semicompatibilists may say “no”)	Yes	Yes
Hard Incompatibilism	No	No	No	No
Revisionism	No	Yes, but only with revision to our self-image	Yes	Yes (or close enough)

# 1

## *Libertarianism*

---

---

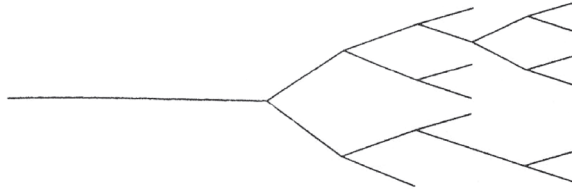
*Robert Kane*

### **1 Determinism and the Garden of Forking Paths**

The problem of free will has arisen in history whenever people have been led to suspect that their actions might be determined or necessitated by factors unknown to them and beyond their control. That is why doctrines of *determinism* or *necessity* have been so important in the history of debates about free will.

Doctrines of determinism have taken many historical forms. People have wondered at various times whether their actions might be determined by Fate or by God, by the laws of physics or the laws of logic, by heredity or environment, by unconscious motives or hidden controllers, psychological or social conditioning, and so on. But there is a core idea running through all historical doctrines of determinism that shows why they are all a threat to free will. All doctrines of determinism – whether they are fatalistic, theological, physical, biological, psychological or social – imply that, given the past and the laws of nature at any given time, there is only one possible future. Whatever happens is therefore inevitable or necessary (it cannot but occur), given the past and the laws.

To see why many persons have believed there is a conflict between free will and determinism, so conceived, consider what free will requires. We believe we have free will when we view ourselves as agents capable of influencing the world in various ways. Open alternatives seem to lie before us. We reason and deliberate among them and choose. We feel (1) it is “up to us” what we choose and how we act; and this means we could have chosen or acted otherwise. As Aristotle said, “when acting is ‘up to us,’ so is not acting.” This “up-to-us-ness” also suggests that (2) the ultimate sources of our actions lie in us and not outside us in factors beyond our control.



**Figure 1** Garden of Forking Paths

To illustrate, suppose Jane has just graduated from law school and she has a choice between joining a law firm in Chicago or a different firm in New York. If Jane believes her choice is a *free* choice (made “of her own free will”), she must believe both options are “open” to her while she is deliberating. She could choose either one. (If she did not believe this, what would be the point of deliberating?) But that means she believes there is more than one possible path into the future available to her and it is “up to her” which of these paths will be taken. Such a picture of an open future with forking paths – a garden of forking paths, it has been called – is essential to our understanding of free will.

This picture of different possible paths into the future is also essential, I believe, to what it means to be a person and to live a human life.

One can see why determinism would threaten this picture. If determinism is true, it seems there would not be more than one possible path into the future available to Jane, but only one. It would not be (1) “up to” her what she chose from an array of alternative possibilities, since only one alternative would be possible. It also seems that, if determinism were true, the (2) sources or origins of her actions would not be in Jane herself but in something else outside her control that determined her choice (such as the decrees of fate, the foreordaining acts of God, her heredity and upbringing or social conditioning).

A second way to illustrate why many people believe there is a conflict between free will and determinism is to reflect on the idea of *responsibility*. Free will is also intimately related to notions of accountability, blameworthiness and praiseworthiness for actions.

Suppose a young man is on trial for an assault and robbery in which his victim was beaten to death. Let us say we attend his trial and listen to the evidence in the courtroom. At first, our thoughts of the young man are filled with anger and resentment. His crime was heinous. But as we listen daily to how he came to have the mean character and perverse motives he did have – a sad story of parental neglect, child abuse, sexual abuse, bad role models –

some of our resentment against the young man is shifted over to the parents and others who abused and mistreated him. We begin to feel angry with them as well as with him. (Note how natural this reaction is.) Yet we aren't quite ready to shift all of the blame away from the young man himself. We wonder whether some residual responsibility may not belong to him. Our questions become: To what extent is *he* responsible for becoming the sort of person he now is? Was it *all* a question of bad parenting, societal neglect, social conditioning, and the like, or did he have any role to play in it?

These are crucial questions about free will and they are questions about what may be called the young man's ultimate responsibility. We know that parenting and society, genetic make-up and upbringing, have an influence on what we become and what we are. But were these influences entirely *determining* or did they "leave anything over" for us to be responsible for? That is what we want to know about the young man. The question of whether he is merely a victim of bad circumstances or has some responsibility for being what he is – the question, that is, of whether he became the person he is *of his own free will* – seems to depend on whether these other factors were or were not *entirely* determining.

Those who are convinced that there is a conflict between free will and determinism, for these and other reasons, are called *incompatibilists* about free will. They believe free will and determinism are incompatible. If incompatibilists also believe that an incompatibilist free will exists, so that determinism is false, they are called *libertarians* about free will.

## 2 Modern Challenges to Libertarian Free Will

I will be defending the libertarian view of free will in this volume. We libertarians typically believe that a free will that is incompatible with determinism is required for us to be truly morally responsible for our actions, so that genuine moral responsibility, as well as free will, is incompatible with determinism. Genuine free will, we believe, could not exist in a world that was *completely* determined by Fate or God, or the laws of physics or logic, or heredity and environment, psychological or social conditioning, and so on. In writings over the past twenty-five years, I have argued that this libertarian view represents the traditional idea of free will that has been in dispute for centuries when philosophers have discussed "the problem of free will and determinism." Moreover, I think this libertarian view is the one many ordinary persons have in mind when they intuitively believe there is some kind of conflict between free will and determinism.

Yet this traditional libertarian conception of free will has been under attack by many modern thinkers, philosophers and scientists alike, who have come

to believe that such an idea of free will, though it may still be held by many ordinary people, is outmoded and incoherent and that it has no place in the modern scientific picture of the world. A goal of this essay is therefore to consider this modern attack on the traditional libertarian view of free will and to ask how, and whether, it can be answered. Much is at stake, it seems to me, in knowing whether we do or do not have a freedom of the will of the ultimate kind that libertarians defend. The modern attack on it has two parts.

*Part 1* The first prong of the modern attack on libertarian free will comes from *compatibilists*, who argue that, despite appearances to the contrary, determinism does not really conflict with free will at all. Compatibilists argue that all the freedoms we recognize and desire in ordinary life – e.g., freedoms from coercion or compulsion, from physical restraint, from addictions and political oppression, for example – are really compatible with determinism. Even if the world should turn out to be entirely deterministic, compatibilists argue, there would still be a big difference between persons who are free from constraints on their freedom of action and will (constraints such as coercion, compulsion, addiction and oppression) and persons who are not free from these constraints; and people would prefer to be free from such constraints on their freedom rather than not, *even in a determined world*. Thus, according to compatibilists, esoteric questions about whether determinism is true or not – in the physical or psychological sciences – are irrelevant to *the freedoms we really care about* in everyday life. All the varieties of free will “worth wanting” (as a modern compatibilist, Daniel Dennett, has put it) do not require the falsity of determinism for us to possess them, as the traditional libertarian view of free will suggests.

This doctrine of *compatibilism* has an ancient lineage. It was held by the Stoics and perhaps also by Aristotle in ancient times, according to many scholars. But compatibilism about free will and determinism has become especially popular in modern times. Influential philosophers of the modern era, such as Thomas Hobbes, John Locke, David Hume and John Stuart Mill, were all compatibilists. They saw compatibilism as a way of reconciling ordinary experience of being free with modern scientific views about the universe and human beings; and compatibilism continues to be popular among philosophers and scientists today for similar reasons, as you will see from later essays of this volume. (John Martin Fischer defends a version of compatibilism, known as *semicompatibilism*, in the second essay of this volume.) If compatibilists are right, we can have *both* free will and determinism; and we need not worry that increasing scientific knowledge about nature and human beings will somehow undermine our ordinary convictions that we are free and responsible agents.



*Part 2* The second prong of the modern attack on libertarian free will goes a step further. Recall that the first prong says that libertarian free will is *unnecessary* because we can have all the freedoms worth wanting, even if determinism should be true. The second prong goes further, arguing that libertarian free will itself is *impossible* or *unintelligible* and has no place in the modern scientific picture of the world. Such an ultimate freedom is not something we could have anyway, say its critics. Those who take this line note that defenders of libertarian free will have often invoked obscure and mysterious forms of agency or causation to defend the libertarian view. In order to explain how free actions can escape the clutches of physical causes and laws of nature (so that free actions will not be determined by physical laws), libertarians have posited transempirical power centers, immaterial egos, noumenal selves outside of 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 modern defenders of libertarianism, such as Immanuel Kant, have argued that we need to believe in libertarian free will to make sense of morality and genuine responsibility, but we can never completely understand such a free will in theoretical and scientific terms.

The problem that provokes this widespread skepticism about the existence of libertarian free will has to do with an ancient dilemma: If free will is not compatible with *determinism*, as libertarians contend, free will does not seem to be compatible with *indeterminism* either (the opposite of determinism). Events that are undetermined, such as quantum jumps in atoms, happen merely by chance. So if free actions were undetermined, as libertarians claim, it seems that they too would happen by chance. But how can chance events be free and responsible actions? 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? Undetermined effects in the brain or body would be unpredictable and impulsive – like the sudden emergence of a thought or the uncontrolled jerking of an arm – quite the opposite of what we take free and responsible actions to be. It seems that undetermined events in the brain or body would occur *spontaneously* and would be more likely to *undermine* our freedom rather than *enhance* it.

This two-pronged modern attack on the traditional libertarian view of free will has had a powerful impact on modern thought. To answer it, libertarians must show (i) that free will really *is* incompatible with *determinism* (call this “The Compatibility Problem”). But they must also show (ii) that a libertarian free will requiring *indeterminism* can be made intelligible and how, if at all, such a free will can be reconciled with modern scientific views of the cosmos and of human beings (call this “The Intelligibility Problem”). I will be

addressing both these problems in this chapter, beginning with the first, or “Compatibility Problem.”

### 3 Is Free Will Incompatible with Determinism?: The Consequence Argument

The popularity of compatibilism among modern philosophers and scientists means that libertarians who believe free will is incompatible with determinism can no longer merely rely on intuitions about “forking paths” into the future to support their view that determinism conflicts with free will (as in section 1). These intuitions must be backed up with arguments that show *why* free will must be incompatible with determinism. To meet this challenge, libertarians have proposed new arguments for incompatibilism in modern philosophy; and we will begin by considering the most widely discussed of these new arguments for the incompatibility of free will and determinism.

This important 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. (From *An Essay on Free Will*, Oxford: Clarendon Press, 1983, p. 16)

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 (it is beyond our control). We can thus spell out this Consequence Argument in the following steps:

- (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.
- (3) There is nothing we can now do to change the past and the laws of nature.
- (4) If determinism is true, our present actions are necessary consequences of the past and the laws of nature. (That is, it *must* be the case that, given the past and the laws of nature, our present actions occur.)
- (5) Therefore, 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 than we actually do (as in the image of forking paths), then no one would have free will.

Defenders of the Consequence Argument, such as van Inwagen, think the first two premises 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 spells out what is implied by determinism. Some philosophers have questioned one or another of the first three steps of this argument. But most criticisms have focused on step 5. Step 5 follows from 3 and 4 by virtue of the following inference: If (3) there is nothing we can now do to change the past and laws of nature and (4) our present actions are necessary consequences of the past and laws, then (5) there is nothing we can now do to change the fact that our present actions occur. This inference is an instance of the following principle:

(TP) If there is nothing anyone can do to change X, and if Y is a necessary consequence of X (if it must be that, if X occurs, Y occurs), then there is nothing anyone can do to change Y.

TP has been called a "Transfer of Powerlessness Principle" for it says in effect that if you are powerless to change something X, and something else Y is necessarily going to occur if X does, then you are also powerless to change Y. This makes sense. If we can't do anything to prevent X from occurring and Y cannot but occur if X does, then how could we do anything to prevent Y from occurring? Consider an example. Suppose the sun is going to explode in AD 2050 and there is nothing anyone can now do to change the fact that the sun will explode in AD 2050. Assume also that necessarily (given the laws of nature), if the sun explodes in AD 2050, all life on earth will end in AD 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 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.

But, despite the initial plausibility of this Transfer of Powerlessness Principle, critics of the Consequence Argument have challenged it. Everything depends, they say, on how you interpret the expression "There is nothing anyone can do to change . . ." Talking about what persons "can" (and "cannot") do is talking about their *powers*; and the notion of power is one of the most difficult in metaphysics, as John Locke pointed out three centuries ago. For example, many *compatibilists* interpret what it means to

say that persons “can” or “have the power” to do things in the following way. They say

“You can (or you have the power to do) something.”

simply means

“If you wanted (or tried) to do it, you would do it.”

I can jump over this fence means I would jump over it, if I wanted to or tried to. If someone challenged my power to do it, the challenger would say “I don’t think you would manage to jump it, *even if* you wanted or tried.”

Now the interesting thing about this compatibilist interpretation of “can” and “power,” is that, if it is correct, the Consequence Argument would fail. For on this interpretation, to say we can now change the past or the laws would mean that

“*If* we now wanted or tried to change the past or the laws, we would change them.”

And this is false. No persons would change the past or the laws of nature, *even if* they wanted or tried to, because no one has the power to do it. But when we turn to ordinary actions like jumping over a fence, things are different. If you can jump over a fence that is in your path, it may well be true that you *would* jump over it, *if* you wanted to or tried, because jumping over fences is something you *are* capable of doing.

In other words, on the analysis of “can” or “power” that many compatibilists favor, the *premises* of the Consequence Argument come out *true* (you would *not* have changed the past or the laws, even if we wanted or tried to, because you are not capable of it). But the *conclusion* of the Consequence Argument comes out *false* (you would have jumped the fence, *if* you wanted or tried to, because jumping fences of this height is something you *are* capable of doing). Since the Consequence Argument would have true premises and a false conclusion on this analysis of “can,” it would be an invalid argument. What has happened to make it fail? The answer is that the transfer principle TP has failed. Your powerlessness to change the past and laws of nature does not *transfer* to your powerlessness to jump the fence. For you are *not* able to change the past and laws, but you *are* able to jump the fence – at least in this compatibilist sense that (“you would do it, *if* you wanted or tried to.”

But why should we accept this “hypothetical” compatibilist account of “can” or “power” (“you would do it, *if* you wanted or tried to”)? Defenders of the Consequence Argument, such as van Inwagen, do not accept this

hypothetical account of “can” or “power”; nor do most libertarians. They would respond to the preceding compatibilist argument as follows:

“So the Consequence Argument fails on your compatibilist analysis of ‘can’ or ‘power.’ But that should not surprise us. For your compatibilist analysis was rigged in the first place to make freedom compatible with determinism. On your analysis, persons can jump the fence even though their doing so here and now is impossible, given the past and the laws of nature. That is not what we libertarians mean by ‘can’ in the Consequence Argument. We mean it is possible that you do it *here and now, given all the facts that presently obtain*. If your analysis allows you to say that persons can do otherwise, even though they can’t change the past and the laws of nature and even though their actions are a necessary consequence of the past and the laws of nature, *then something must be wrong with your compatibilist analysis*. What use is a power or ability to do something, if it cannot be *exercised* in the existing circumstances here and now? To us libertarians, the premises and rules of the Consequence Argument are far more plausible than any compatibilist analysis of ‘can.’”

At this point, arguments over the Consequence Argument tend to reach an impasse. Incompatibilist defenders of the argument claim that compatibilist critics are begging the question by interpreting “can” in the Consequence Argument in a way that is compatible with determinism. But compatibilists respond by saying that defenders of the Consequence Argument are begging the question themselves by assuming that “can” in the argument has an *incompatibilist* meaning rather than a compatibilist one.

#### 4 Ultimate Responsibility

As a result of this impasse, philosophical debates have multiplied about just what “can” and “power” (and related expressions, such as “could have done otherwise”) really mean. We cannot follow all these complex debates here. But I do not think it matters. For I believe disagreements over the meaning of “can” and “power” are symptoms of a deeper problem in discussions about free will and determinism. The problem is that focusing on “alternative possibilities” (or “forking paths” into the future) or the “power to do otherwise” *alone*, as the Consequence Argument does, is *too thin a basis* on which to rest the case for the incompatibility of free will and determinism. One must look beyond debates about “can,” “power,” “ability,” and “could have done otherwise” to make the case for the incompatibility of free will and determinism.

Fortunately, there is another place to look for reasons why free will might conflict with determinism. Recall that in section 1, I suggested that there were *two* reasons why people thought determinism must rule out free will.

One was the requirement of (1) alternative possibilities we have been considering: Free will seems to require that *open alternatives* or *alternative possibilities* lie before us – a garden of forking paths – and it is “up to us” which of these alternatives we choose. (Call this condition “AP” for “alternative possibilities”). But there was a second condition mentioned that has also historically fueled incompatibilist intuitions: (2) Free will also seems to require that the *sources* or *origins* of our actions lie “in us” rather than in something else (such as the decrees of fate, the foreordaining acts of God, or antecedent causes and laws of nature) outside us and beyond our control.

I call this second requirement for free will the condition of Ultimate Responsibility (or UR, for short); and I think it is even more important to free will debates than AP, or alternative possibilities. The basic idea of UR is this: *To be ultimately responsible for an action, an agent must be responsible for anything that is a sufficient cause or motive for the action's occurring.* If, for example, a choice issues from, and can be sufficiently explained by, an agent's character and motives (together with background conditions), then to be *ultimately* responsible for the choice, the agent must be at least in part responsible by virtue of choices or actions voluntarily performed in the past for having the character and motives he or she now has. Compare Aristotle's claim that if a man is responsible for the wicked acts that flow from his character, he must at some time in the past have been responsible for forming the wicked character from which these acts flow.

This condition of Ultimate Responsibility, or UR, does not require that we could have done otherwise (AP) for *every* act done of our own free wills. But it does require that we could have done otherwise with respect to *some* acts in our past life histories by which we formed our present characters. I call these earlier acts by which we formed our present characters “self-forming actions,” or SFAs.

To see why such self-forming acts are important for free will, consider a well-known example about Martin Luther offered by Daniel Dennett. When Martin Luther finally broke with the Church in Rome, initiating the Protestant Reformation, he said “Here I stand, I can do no other.” Now Dennett asks us to suppose that at the moment Luther made this stand, he was literally right. Given his character and motives, Luther *could* not then and there *have done otherwise*. Does this mean Luther was not morally responsible, not subject to praise or blame, for his act, or that he was not acting of his own free will? Dennett says “not at all.” In saying “I can do no other,” Luther was not disowning responsibility for his act, according to Dennett, but taking full responsibility for acting of his own free will. So the ability to do otherwise (“could have done otherwise”) or AP, says Dennett, is not required for moral responsibility or free will.

Now Dennett is a compatibilist, as noted earlier, and he is using this Luther example to defend compatibilism of free will and determinism by

suggesting that free will and moral responsibility do not even require the power to do otherwise or alternative possibilities (AP). Note that, if this were true, the Consequence Argument would be undermined. We would not have to get into complex debates about what “could have done otherwise” means, since free will and moral responsibility would not require alternative possibilities (AP) or “could have done otherwise” in the first place.

But, now, if we look at Dennett’s Luther example from the point of view of the condition of Ultimate Responsibility or UR, rather than simply in terms of AP, there is an answer that can be given to Dennett. We can grant that Luther could have been responsible for this act, even though he could *not* have done otherwise *then and there* and even if his act was determined. But this would be so, if UR is required, only to the extent that Luther was responsible for his present motives and character by virtue of some *earlier* struggles and self-forming actions (SFAs) that brought him to this point in his life where he could do no other. Those who know Luther’s biography know the inner struggles and turmoil he endured getting to that point in his life. Often we act from a will already formed, but it is “our own free will” by virtue of the fact that *we* formed it by other choices or actions in the past (SFAs) for which we *could* have done otherwise. If this were not so, *there is nothing we could have ever done to make ourselves different than we are* – a consequence, I believe, that is incompatible with our being (at least to some degree) ultimately responsible (UR) for what we are. So SFAs are only a subset of those acts in life for which we are ultimately responsible and which are done “of our own free will.” But if *none* of the acts in our lifetimes were self-forming in this way, we would not be *ultimately* responsible for anything we did.

If the case for incompatibility of free will and determinism cannot be made by reference to AP alone, it can be made if UR is added. So, I suggest, the often-neglected condition of ultimate responsibility or UR should be moved to center stage in free will debates. If agents must be responsible to some degree for anything that is a *sufficient cause* or *motive* for their actions (as UR requires), then an impossible infinite regress of past actions would be required unless some actions in the agent’s life history (SFAs) did not have either sufficient causes or motives (and hence were undetermined). Therein lies the connection between UR and determinism. If we must have formed our present wills (our characters and motives) by earlier voluntary choices or actions, then UR would require that if any of these earlier choices or actions *also* had sufficient causes or motives when we performed *them*, then we must have also been responsible for those earlier sufficient causes or motives by virtue of forming them by *still earlier* voluntary choices or actions, and so on backwards indefinitely into our past. Eventually we would come to infancy or to a time before our birth when we could not have formed our own wills.

The only way to stop this regress is to suppose that *some* acts in our life histories must lack *sufficient* causes altogether, and hence must be undetermined, if we are to be the ultimate sources or grounds of, and hence ultimately responsible for, our own wills. These regress-stopping acts would be the “self-forming acts” or SFAs that are required by UR sometime in our lives, if we are to have free will. Note, as a result, that UR makes explicit something that is often hidden in free will debates, namely that *free will* – as opposed to mere *freedom of action* – is about the forming and shaping of character and motives which are the *sources* or *origins* of praiseworthy or blameworthy, virtuous or vicious, actions. *Free will* (in contrast to mere *free action*) is about *self-formation*. If persons are responsible for the wicked (or noble, shameful, heroic, generous, treacherous, kind or cruel) acts that flow from their wills (characters and motives), they must at some point be responsible for forming the *wills* from which these acts flow.

## 5 Ultimate Responsibility and Alternative Possibilities

Another thing to note about this argument for the incompatibility of free will and determinism from UR is that – unlike the Consequence Argument – the argument from UR does not mention the condition of *alternative possibilities* or AP at all. The argument from UR says that, if agents must be responsible to some degree for anything that is a *sufficient cause or motive* for their actions (as UR requires), then an impossible infinite regress of past actions would be required, *unless* some actions in the agent’s life history (SFAs) did not have either sufficient causes or motives and hence were undetermined. The argument from UR thus focuses on the sources or origins of what we actually do rather than on the power to do otherwise.

When one argues about the incompatibility of free will and determinism from alternative possibilities or AP (as in the Consequence Argument), the focus is on notions of “necessity,” “possibility,” “power,” “ability,” “can,” and “could have done otherwise.” By contrast, the argument from UR focuses on a different set of concerns about the “sources,” “grounds,” “reasons,” and “explanations” of our wills, characters, and purposes. Where did our motives and purposes come from, who produced them, who is responsible for them? Was it *we* ourselves who are responsible for forming our characters and purposes, or someone or something else – God, fate, heredity and environment, nature or upbringing, society or culture, behavioral engineers or hidden controllers? Therein lies the core of the traditional problem of free will.

But does this mean that alternative possibilities or AP have nothing to do with free will? It might seem so, if one can argue directly for the incompatibility of free will and determinism from UR without mentioning alternative