

Freedom and Determinism

edited by Joseph Keim Campbell, Michael O'Rourke, and David Shier

**A Bradford Book
The MIT Press
Cambridge, Massachusetts
London, England**

© 2004 Massachusetts Institute of Technology

All rights reserved. No part of this book may be reproduced in any form by any electronic or mechanical means (including photocopying, recording, or information storage and retrieval) without permission in writing from the publisher.

This book was set in Stone Serif and Stone Sans by Interactive Composition Corporation and was printed and bound in the United States of America.

Library of Congress Cataloging-in-Publication Data

Freedom and determinism / edited by Joseph Keim Campbell, Michael O'Rourke, and David Shier.

p. cm.—(Topics in contemporary philosophy)

“A Bradford book.”

Includes bibliographical references and index.

ISBN 0-262-03319-4 (hc. : alk. paper)—ISBN 0-262-53257-3 (pbk. : alk. paper)

1. Free will and determinism. I. Campbell, Joseph Keim, 1958– . II. O'Rourke, Michael, 1963– . III. Shier, David, 1958– . IV. Series.

BJ1461.F755 2004

123'.5—dc22

2003066491

10 9 8 7 6 5 4 3 2 1

10 Van Inwagen on Free Will

Peter van Inwagen

I can remember very clearly the first time van Inwagen encountered the problem of free will. In the autumn of 1965 he was talking with a fellow graduate student at the University of Rochester, one Myles Brand, and made some remark that presupposed the incompatibility of free will and determinism. Brand told him—second-year graduate student to first-year graduate student—that most philosophers believed that free will and determinism were compatible, and outlined some of the currently popular arguments for that position. As Athena from the head of Zeus, the argument that van Inwagen was to publish ten years later in “The Incompatibility of Free Will and Determinism” (1975) sprang from his head pretty much full-grown—although it made its entrance into the world by way of his ever-active mouth and not by Athena’s rather more unorthodox route.

The argument had its roots in the following reflections. If free will and determinism coexist, then someone is able to do something not contained in that one possible future that is consistent both with the past and the laws of nature. Suppose that Alice, an inhabitant of a deterministic world, is able to do something she is not in fact going to do; suppose, to be specific, that although she is going to remain a prisoner, she is able to escape from her prison. Her ability to escape can be looked upon as an ability to divert the river of coming events into a channel through which it is not in fact going to flow; to realize, that is, or to cause to be actual, a possible future that is not the future that lies before her, to cause to be actual one of those possible futures in which she escapes. And what would these possible but nonactual futures be like? Let’s say that the past *of* a possible but nonactual future is that possible past that would be the actual past if that future were the actual future. Any given future in which Alice escapes must either be a future whose past is the actual past and which is discontinuous with that past, or a

future that is continuous with its past but whose past is not the actual past (a past different from the actual past all the way back to the Big Bang), or, finally, a future in which the world is governed by laws of nature that are not the actual laws. But Alice can bring about futures of none of these sorts. If it is insufficiently evident as it stands that she can bring about none of these futures, here is an argument. Let ‘Clio’ be a proper name for the actual past—thus, when I imagine, as I am about to, Alice using the name ‘Clio’ in another possible world, I imagine her referring to the past as it is in our world, not the past as it is in hers. Similarly, let ‘Nomos’ be a proper name for the actual laws of nature (which, remember, we are assuming to be deterministic). It would seem that if Alice is able to escape, she must be able in some sense to cause the actuality of or bring about or realize a future in which she could say, and in so saying speak truly, one of the following three things:

- There has been a causal break; the present state of affairs is not continuous with the past.
- The present is continuous with the past, but that past is not Clio; it is some other past (a past different from Clio all the way back to the Big Bang).
- The laws of nature are not Nomos, but some other set of laws.

And, obviously, Alice is not able to get herself into a future in which she can say any of these things and be right.

When van Inwagen got around to writing down the argument that had occurred to him in his conversation with Brand—he first did this in a doctoral thesis he wrote under the supervision of Richard Taylor (*de jure*) and Keith Lehrer (*de facto*)—the argument he wrote down did not *look* much like the argument I have just set out. (It looked very much like the argument he would later publish in “The Incompatibility of Free Will and Determinism.”) Nevertheless, the central idea of both arguments was the same, and they no doubt stand or fall together.

When van Inwagen had got his first academic job and was trying to publish this argument, he found it extraordinarily hard to do so. The reason was simple: the conclusion of the argument was known to be false. The unanimous position of the referees whose reports were enclosed with the rejection letters could have been expressed in the following words, which I take from an essay by Donald Davidson, words as well known as they are ungrammatical:

I shall not be directly concerned with such arguments [i.e., arguments for the incompatibility of free will and determinism], since I know of none that is more than superficially

plausible. Hobbes, Locke, Hume, Moore, Schlick, Ayer, Stevenson, and a host of others have done what can be done, or ought ever to have been needed, to remove the confusions that can make determinism seem to oppose freedom. (1973, 137)

It's not that no one was an incompatibilist in those days, of course. The volume that contained the essay from which my quotation from Davidson is taken also contained David Wiggins's "Toward a Responsible Libertarianism." Roderick M. Chisholm was an incompatibilist, as were Richard Taylor, Carl Ginet, Elizabeth Anscombe, and Peter Geach. Still, it can hardly be denied that incompatibilists were thin on the ground in the sixties and early seventies.

Eventually, however, van Inwagen was able to publish two papers in which he argued for the incompatibility of free will and determinism. (He has always suspected that "The Incompatibility of Free Will and Determinism," which had been rejected by many journals, was accepted by *Philosophical Studies* only because Sellars chose Keith Lehrer as its referee, but he has never known whether this suspicion was correct.) His other paper on the subject, "A Formal Approach to the Problem of Free Will and Determinism" (1974), had, in respect of publication, the advantage of making use of the then new and exciting apparatus of possible worlds. (This paper was strongly influenced by Montague's "Deterministic Theories" (1974); the impetus for its writing was a suggestion of Rolf Eberle's.)

Van Inwagen summed up his thought on free will in his book *An Essay on Free Will* (1983), and has pretty much avoided learning anything about the problem since—other than by sitting about and thinking it over. (The publication of this book by Oxford University Press was due to the good offices of Tony Kenny and Derek Parfit, neither of whom could have had any sympathy whatever with its content, and van Inwagen has always been grateful to them for their generosity and has tried to imitate it. He has done his best to learn from them the hard lesson that a philosophical book that he regards as thoroughly wrongheaded can nevertheless be a good book.) Van Inwagen likes to think that this book bears a significant share of the responsibility for the fact that incompatibilists are now much more common than they were thirty or forty years ago. In a paper that he read at a conference in the early nineties, van Inwagen made a remark to the effect that compatibilism was the standard view among philosophers. Michael Slote, who was in the audience, said that he thought that, on the contrary, incompatibilism had become the standard view, or at least the majority view.

A few years later, van Inwagen asked Ted Warfield whether he thought that was right. Warfield, who comes as close as is humanly possible to knowing what every analytical philosopher thinks about anything, replied that he thought that the majority of analytical philosophers who had actually worked on the free-will problem were incompatibilists, and that the majority of analytical philosophers (full stop) were compatibilists.

There was one passage in *An Essay on Free Will* that, after the publication of the book, van Inwagen became more and more worried about.¹ (These worries were aggravated by pointed questions from Alex Rosenberg, and also, curiously enough, by a science-fiction novel by Larry Niven called *Protector*—a novel about which van Inwagen is ready to tell you considerably more than you want to know.) Van Inwagen had said in that passage that although no one was able to render a *physical* law false, it seemed that if there were *psychological* laws, and if we had free will, we had to be able to render *these* laws false. But that raises the question: if one has it within one's power to render some proposition false, in what sense can that proposition be a law? As a sort of schematic example of a psychological law, van Inwagen proposed, "No one who has received moral training of type A in early childhood ever spreads lying rumors about his professional colleagues in order to advance his career." He imagined someone asking, "Why does the pattern of behavior described in this statement occur if people don't *have* to conform to it?" He answered this question as follows: "Perhaps it is just the people who have received moral training of type A in early childhood who *see the point* in not spreading lying rumors about their colleagues."

He gradually came to see, or to think he saw, that this response to the difficulty was facile, and that the difficulty he had his finger on was broader and deeper than the original puzzle about psychological laws. He gradually came to the conclusion that if one was faced with the necessity of doing either A or B, and that if one saw every reason to do A and no reason whatever to do B, then one would simply not be able to do B. From this conclusion it was no great leap to the slightly stronger conclusion that, if one was faced with a choice between A and B, and one was aware of considerations that could be brought in support of both alternatives, and if the considerations that supported A seemed to one clearly and decisively to outweigh the considerations that supported B, then one would simply not be able to do B. Van Inwagen defended, in "When Is the Will Free?" (1989), the thesis that the general principles about ability that lead philosophers to

incompatibilism should lead anyone who accepts them to accept these conclusions as well. And he went on to argue that, since occasions that call for serious deliberation—occasions, that is, on which one is choosing between alternatives and it does not seem to one that (once all the purely factual questions have been settled) the reasons that favor either alternative are clearly the stronger—at best only a small proportion of the occasions on which we make a choice are occasions on which we make a free choice. ('At best' because there may be no free will at all; perhaps determinism is true, or perhaps—as Broad believed—free action is incompatible both with determinism and indeterminism and is therefore a self-contradictory idea.) Van Inwagen concluded that no action is free unless it is the outcome of deliberation in which one considers reasons that support that act, reasons that support various alternative acts, and in the course of which one finds no obvious answer to the question, 'Which set of reasons should prevail?' To take one example among many different sorts of possible example of the consequences of this position, if you answer the telephone 'automatically', if you answer the telephone without so much as considering the question whether you should answer it, your act is not a free act: you could not have done otherwise than answer the telephone; you were not *able* to let it ring till it fell silent; it was *not within your power* not to raise the receiver.

After presenting arguments for this thesis, van Inwagen went on to attempt to show that it does not, or does not obviously, have a certain untoward consequence that it might be thought to have. He contended that from the premise that at best a very small proportion of our acts are free acts, the conclusion does not follow that only a very small proportion of our ascriptions of moral responsibility are correct. For, he maintained, although there is an inseverable connection between free will and moral responsibility, this connection, inseverable though it be, can be stretched exceedingly fine. An example will illustrate his point. Suppose a man is driving drunk and that a pedestrian suddenly looms before him. He attempts to swerve, but too late: he hits and kills the pedestrian because his reflexes are impaired by alcohol. Compare his case with the case of a sober, able, and alert driver whose car strikes and kills a pedestrian in circumstances in which swerving in time to avoid the pedestrian would have required a reaction time smaller than that allowed by the speed of propagation of human neural impulses. In neither case was the driver able to avoid hitting the pedestrian who suddenly loomed before him, but when we consider the former case, we hold the

driver morally responsible for the pedestrian's death, and when we consider the latter case we do not.

The relevant difference is, of course, that the man whose reflexes were impaired by drink, was, so to speak, *able* to avoid being unable to avoid hitting the pedestrian, and the sober and alert driver was *unable* to avoid being unable to avoid hitting the pedestrian. At the moment he first saw the pedestrian, the drunk driver was unable to avoid hitting him, but he had earlier been able, or so we should suppose if we were judge or jury, to avoid driving with the impaired reflexes that were the cause of his fatal inability at the time of the accident. Van Inwagen suggested that this sort of case could serve as a model for the relation between ability and responsibility. Here is a second case, a case in which the inseverable connection between ability and responsibility, though it remains unsevered, as inseverable connections do, is stretched considerably further than it is in most philosophical examples concerning moral responsibility and the ability to do otherwise. Consider a man who is, in middle age, a corrupt politician and is, owing to his corrupted nature, unable to refuse bribes when he believes there is no significant likelihood of the bribery coming to light. That is how he is, but how did he get that way? Suppose the answer is this: as a young man, he made a certain series of free choices, choices preceded by genuine deliberation, which collectively had the effect of establishing him in settled and unbreakable habits of venality. Van Inwagen argued—guided, I suppose, by Aristotle—that this politician can properly be held morally responsible for the baleful effects on the public welfare of the informal services he renders to his political cronies in return for money. And this despite the fact that he is unable, in middle age, to reject the bribes he is offered. He can properly be held responsible for, say, the deaths of the four children in the fire in the building that wasn't up to code, because he could, as a young man, have avoided becoming the sort of man who would be unable to resist the bribe offered by the slumlord who owned the building.

Several philosophers have disputed van Inwagen's conclusion that the principles that lead philosophers to incompatibilism entail that free acts, if they exist at all, are extremely rare, but van Inwagen has never been able to see any force in their arguments. Although he has published answers to them (1994), he is of the opinion that no answers were needed; that his original arguments were untouched by the arguments of his critics. One philosopher, who generally disagrees with van Inwagen about free will,

Dan Dennett, agrees with van Inwagen that these arguments are unanswerable. As Dan put it, referring to “When Is the Will Free?”, “Thank you, Peter, for the lovely *reductio* of incompatibilism.”

Now van Inwagen’s arguments for this conclusion, whether they are good or bad, presuppose that there is an inseverable connection between moral responsibility and the power to do otherwise, however flexible this connection may be. The inseverable connection is this: if one is morally responsible for anything, it follows logically that one has had a free choice about something. But Harry Frankfurt has presented a famous argument that some have taken to refute this thesis.² A significant proportion of van Inwagen’s work on free will has been devoted to Frankfurt’s argument. And Frankfurt’s argument is important. If it is indeed true that one might be morally responsible for various things, despite one’s *never* having been able to do otherwise than one has done, then the problem of free will loses much of its interest—for the simple reason that most people would find the thesis that we lack free will much less unappealing if this thesis could be shown not to entail that we can never be held to moral account for anything.

I have said that Frankfurt’s argument has been taken by some to show that it is possible for one to be morally responsible for something even though one has never been able to do otherwise. The actual conclusion of Frankfurt’s argument, however, is this: the so-called principle of alternative possibilities is false, or at least not a necessary truth. (I’ll call it the principle of *alternative* possibilities. Frankfurt has recently presented an ill-advised defense of his use of the adjective ‘alternate’ in his name for the principle. It is, I concede, uncharitable of me to mention this. I’ll attempt to atone for my lapse by very charitably saying nothing further about it.) This is the principle of alternative possibilities:

A person is morally responsible for what he has done only if he could have done otherwise.

Van Inwagen has always thought that Frankfurt’s argument—which, of course, consists in the presentation of a certain sort of counterexample to the principle of alternative possibilities—has a great deal of force and has never been shown conclusively to be mistaken. His position has never been that Frankfurt’s proposed counterexamples to the principle of alternative possibilities fail; his position has been rather that even if these counterexamples

succeed, even if the principle of alternative possibilities is false, the existence of moral responsibility nonetheless requires the existence of free will.

I have used the qualified phrases ‘has a great deal of force’ and ‘has never been shown conclusively to be mistaken’ because, although van Inwagen is inclined to think that Frankfurt’s counterexamples show that the principle of alternative possibilities is false if it is meaningful at all, he’s also inclined to think that it’s meaningless. I’ll briefly say something about why he’s inclined to think it’s meaningless. If the principle of alternative possibilities is meaningful, the following must be one of the particular statements the general principle endorses:

Bill is morally responsible for lying under oath only if he could have done otherwise.

Van Inwagen has never been able to convince himself that he understands sentences like ‘Bill is morally responsible for lying under oath’. It has always seemed reasonably plain to him that what one is morally responsible for is not one’s acts but the *consequences* of one’s acts, or, more exactly, *certain* of the consequences of one’s acts—for no one would suppose that one could be responsible for *all* the consequences of one’s acts. (When I say ‘it has always seemed plain to him’, I mean that it has seemed plain to him since he first encountered the idea that what a person is morally responsible for is the consequences of his acts and not the acts themselves. Van Inwagen is sure that he first encountered this idea in something written by P. H. Nowell-Smith; he cannot now remember where it was that Nowell-Smith made this valuable point.) And the consequences of one’s acts, it would seem, are members of the same ontological category or categories as the consequences of anything that takes place within the causal and temporal order; whatever ontological category one thinks the consequences of a person’s acts should be assigned to, one should assign them to the same category or categories as the consequences of an earthquake or a scientific discovery or a rise in the prime lending rate. There would seem to be two serious candidates for this categorial office: ‘concrete event’ (for example, Caesar’s death), on the one hand, and something proposition-like on the other, ‘fact’, perhaps, or ‘state of affairs’ (for example, the fact that Caesar disregarded the soothsayer’s warning, or Caesar’s having chosen to believe that Brutus’ strong republican sentiments would never overcome his friendship with and personal loyalty to Caesar).

Now if what one is responsible for is certain of the consequences of one's acts (and if, for example, one's telling a lie and the fact that one has lied do not count as 'consequences' of themselves), then it is doubtful whether 'Bill is responsible for lying under oath' makes sense, and it is therefore doubtful whether the principle of alternative possibilities makes sense, doubtful whether the sentence that formulates it means anything. Van Inwagen, as I have said, has long doubted whether the principle of alternative possibilities does make sense. But he has also been fairly sure that if he's wrong about this and the sentence 'A person is morally responsible for what he has done only if he could have done otherwise' does express some proposition, the proposition it expresses is false, and that Frankfurt's counterexamples show this.

To recapitulate, van Inwagen thinks that (a) the principle of alternative possibilities is either nonsensical or false, and that (b) moral responsibility nevertheless requires free will—that if anyone is morally responsible for anything, there must be something that person had a free choice about.

He has defended the latter thesis by presenting other principles than the principle of alternative possibilities that have the consequence that moral responsibility requires free will, and presenting arguments designed to show that these other principles seem to be true and cannot be refuted by counterexamples in the style of Frankfurt's counterexamples to the principle of alternative possibilities.³ One of these principles is

A person is morally responsible for a certain state of affairs only if that state of affairs obtains and he could have prevented it from obtaining.

(Or this was van Inwagen's original formulation of one of these principles. He later came to think that the principle needed to be revised. The revised principle would look something like this:

A person is morally responsible for a certain state of affairs only if that state of affairs obtains and there was a time at which he could so have acted that that state of affairs not obtain.

The states of affairs quantified over in this principle are 'proposition-like': the state of affairs 'Caesar's having been murdered' obtains because certain conspirators stabbed Caesar to death in Rome in 44 B.C.; but it, that very same state of affairs, could have obtained because Cleopatra poisoned him in Alexandria in 48—just as the proposition that Caesar was murdered is

true because certain conspirators stabbed Caesar to death in Rome in 44 B.C. and could have been true because Cleopatra poisoned him in Alexandria in 48. This state of affairs may thus be contrasted with the concrete event *the murder of Caesar*, which would not have occurred if Cleopatra had poisoned Caesar in Alexandria in 48, although, in that case, there would have been a concrete event, which does not in fact exist, that would have been denoted by the words 'the murder of Caesar'. (He has also endorsed a principle about concrete events that corresponds to or parallels the above principle about states of affairs.)

Van Inwagen has never seen any need to rethink the position he took concerning this principle when he first formulated it in the late seventies (with this minor qualification: as I have said, he has come to prefer a revised version of the principle), to wit that it is extremely plausible that it entails that moral responsibility requires free will, and that it cannot be refuted by any adaptation of the counterexamples Frankfurt brought against the principle of alternative possibilities. This last point has been disputed, but van Inwagen has never been able to see any merit in the arguments by means of which the point has been disputed. He has responded to them (1997 and 1999), but, as far as he is concerned, his original arguments for this position are the only answer to these counter-arguments that was really needed.

This brings us down to fairly recent times. (Or it does if we ignore two minor essays of the middle eighties. In one of these essays [1985a], van Inwagen defended the position that the 'conditional-analysis' argument for the compatibility of free will and determinism, and what he has dubbed the *Mind* Argument for the incompatibility of free will and indeterminism—the latter argument is named for the journal, not for the human intellectual faculty—are inconsistent; inconsistent in this sense: from the soundness of either, the unsoundness of the other follows. In the other essay [1985b], he proposed a conditional-analysis argument for the compatibility of free will and indeterminism; this argument was, admittedly, a rather absurd argument; his point was that the unbiased inquirer should see that the conditional analysis argument for the compatibility of free will and determinism was equally absurd, and absurd for an exactly parallel reason.) Van Inwagen has thought little about free will in the last ten years, but he has thought enough about the topic to have changed his mind on one point and to have become increasingly insistent on another.⁴

The point on which he has changed his mind is this: he now thinks that Rule β , as he presented it in *An Essay on Free Will* and other places, is wrong. This rule of inference was stated as follows:

p, and no one has, or ever had, any choice about that. If *p* then *q*, and no one has, or ever had, any choice about that. Hence, *q*, and no one has, or ever had, any choice about that.

Van Inwagen had thought that this rule was obviously valid, had made use of it in formulating one version of his argument for the incompatibility of free will and determinism (his favorite version, the version inspired by Carl Ginet's "Might We Have No Choice?" [1966]), had said that, despite its seeming obviousness, it was nevertheless the weakest link in the chain of reasoning that led from the assumption of determinism to the conclusion that no one is ever able to do otherwise. And, although this principle is used in only one of the three formal versions of the argument he has presented, van Inwagen is on record as saying that, in his opinion, if Rule β should turn out to be invalid, it would almost certainly be the case that the two versions of the argument that do not involve the concept 'having a choice' would also turn out to be invalid.⁵

Imagine, then, his embarrassment when Tom McKay and David Johnson (1996) presented a counterexample to Rule β —a 'beta blocker', to use the term that some wag (was it John Martin Fischer?) coined for counterexamples to Rule β . (Actually, McKay and Johnson presented a counterexample to a different rule of inference, but since the invalidity of that other rule implies the invalidity of β , they have for all practical purposes presented a counterexample to β . It is not, by the way, entirely clear why they proceeded in this indirect fashion, since the idea behind their counterexample to the other rule of inference can easily be adapted to produce a counterexample to β .⁶)

Van Inwagen (2000, 8–9) has presented a revised rule of inference that is immune to McKay–Johnson-style counterexamples. Finch and Warfield (1998) have proposed a different sort of revision that is also immune to these counterexamples, and McKay and Johnson themselves suggested some revisions that are, I think, workable. Unfortunately, these revised rules, although they are far from implausible, lack the 'luminous evidence' that was a striking, albeit illusory, property of the original Rule β . And this fact can very properly prompt the compatibilist to ask a pointed question

along the following lines: “If the apparent intuitive obviousness of the original Rule β turned out to be an illusion—since the rule has in fact turned out to be invalid—how much confidence should we have in the revised versions? Nemesis, in the form of a counterexample, was all along lying in wait for Rule β ; why should we not take very seriously the epistemic possibility that in some dark corner of logical space, cousins of this Nemesis patiently await their appointments with the revised rules?” This pointed question is pointed indeed. Van Inwagen admits that he has no good answer to this question, and that, in consequence, although he *accepts* the revised rule, he assigns its validity a rather lower subjective probability than the near certainty he once so confidently assigned to the validity of the original β . He does, however, withdraw his assent to a thesis I mentioned a moment ago, the thesis that if Rule β should turn out to be invalid, this would mean that there was almost certainly something wrong with the other arguments he has given for the incompatibility of free will and determinism. He withdraws his assent because the counterexamples that have shown the invalidity of Rule β , the McKay–Johnson counterexamples, depend on some unexpected properties of an English phrase—“has a choice about”—that played a key role in his formulation of the original principle. He is inclined to think that the ‘general idea’ behind Rule β was sound, and that its invalidity stemmed from the fact that certain features of the English phrase unfit it for the task he assigned to it. When he made the statement from which he has now withdrawn his assent, he was assuming that if Rule β were shown to be invalid, this would be because someone had shown that the general idea behind the principle was fundamentally defective. But the revised versions of the rule appeal to this same general idea; they are merely (it is hoped) adequate implementations of this idea, implementations from which a technical defect has been removed.

I have said that in recent times, van Inwagen has changed his mind about one point and has become increasingly insistent on another. I have discussed the point on which he has changed his mind. The point on which he has become increasing insistent is this: free will is a mystery, a ground-floor, first-water, Colin McGinn–style philosophical mystery. Free will is a mystery because, although it obviously exists—of *course* we sometimes confront a choice between A and B and are, while we are trying to decide whether to do A or to do B, able to do A and able to do B—it seems to be incompatible with both determinism and indeterminism, and thus seems to

be impossible. When he says that free will seems to be incompatible both with determinism and indeterminism, van Inwagen means that there are good arguments for the incompatibility of free will and determinism and good arguments for incompatibility of free will and indeterminism, and that no one has ever identified a very plausible candidate for the flaw in any of the arguments in either class. Van Inwagen, of course, believes that the arguments he has given for the incompatibility of free will and determinism contain *no* flaws—or, at the worst, contain minor, technical flaws that could be repaired without altering their essential points—and that there is some flaw, or are some flaws, in the familiar arguments for the incompatibility of free will and indeterminism. But as to the latter class of arguments—well, he’s damned if he knows what the flaws in them might be. He simply hasn’t a clue. If you ask him how free will is possible, he can say only, “That’s a mystery.” This is, of course, a situation that a philosopher must regard as unsatisfactory, but van Inwagen has two consolations. First, he thinks that if he believed that an undetermined act could not be a free act, he’d be forced to believe something even more mysterious than what he now believes. He’d be forced to believe that at least one of the following three propositions was false:

- Free will is incompatible with determinism
- Some people are morally responsible for various things, but no one has free will
- People are sometimes morally responsible for various things.

And the falsity of any of these propositions would be, in his view, an even greater mystery than the falsity of the proposition that free will is compatible with indeterminism. Second, there are *lots* of philosophical mysteries, most of which have nothing to do with free will. I might cite one or more of the great philosophical mysteries, such as ‘temporal passage’ or human consciousness. But I will mention instead two minor mysteries that are no less mysterious for being minor. I will mention two mysteries that carry much less emotional weight than the mystery of free will, or are, as one might say, unconnected with anything we care deeply about—unlike free will and time and consciousness, which are connected with many things that we care deeply about: the mystery of vagueness and the mystery of the Liar. It is obviously true that a currently living American male who is seven feet four inches tall is tall, and is tall without qualification (i.e., is not

simply tall to degree 0.99972 or something like that). It is obviously true that the president-elect, falsely believing himself to have lost the election, might address an audience and say, bitterly and sincerely, "Everything the president-elect tells anyone today is false." I do not think that either of these obvious truths is free from apparent paradox—although, of course, no paradox can be more than apparent, for no truth can *really* imply a self-contradictory statement or even a false statement. And yet there are very good arguments for the conclusion that each of these obvious truths entails various statements that are self-contradictory or at least obviously false. And I cannot find any flaw in these arguments, though, of course, there must *be* flaws in them. Van Inwagen's expectation is that some premise (or more than one) in the several arguments for the incompatibility of free will and indeterminism is false but seems to us to be an obvious truth—as was once the case with the proposition that there are twice as many natural numbers as there are even numbers, the Galilean law of the addition of velocities, the argument that, since space is unbounded, it must be infinite, or the unrestricted comprehension principle in set theory. But if van Inwagen is wrong about this, the following is *certainly* the case: *some* proposition (or maybe there is more than one) about matters relating to free will, determinism, and moral responsibility that seems to us to be obviously true is false. Perhaps we shall one day discover what proposition this is, or what propositions these are, and it, or they, will come to have no more appeal to philosophers than the proposition that there are twice as many natural numbers as there are even numbers (which Galileo thought evident, despite his awareness that the two classes of numbers could be put into one-one correspondence) has for us. Or perhaps, as Chomsky (1988, 151–152) and McGinn (1993) have suggested, some evolutionarily contingent feature of the design of the human intellect renders it biologically impossible for us to think our way through the free will problem to a satisfactory conclusion. I will remark, because it makes a nice link with my next topic, that van Inwagen has strong, I might almost say fraternal, feelings for Chomsky when he considers Chomsky's position on free will and determinism. Chomsky thinks that free will must be compatible with determinism, and that, nevertheless, it is a mystery how this could be. Van Inwagen thinks that free will must be compatible with indeterminism, and that, nevertheless, it is a mystery how this could be. Van Inwagen's feeling is that Chomsky, although he is mistaken about a particular point (an

important point, to be sure), appreciates the depth and difficulty of the free will problem, while the majority of van Inwagen's fellow incompatibilists, although they are right on a particular point (and a very important point it is), do not really appreciate the depth and difficulty of the problem.

I want to close by explaining why van Inwagen thinks one important group of incompatibilists, those who appeal to what is called agent causation, do not appreciate the depth and difficulty of the problem of free will. Many philosophers would agree with this judgment for the simple reason that they think that the concept of agent causation is incoherent, or think that agent causation is metaphysically impossible. Van Inwagen is inclined to agree with them (although he has no firm opinion on this question), but he has lately stressed a different point. It is this: suppose there is nothing conceptually or metaphysically impossible about agent causation; suppose in fact that agent causation is a real phenomenon and that an episode of agent causation figures among the antecedents of every voluntary movement of a human hand or limb or vocal apparatus. Van Inwagen's position is that even if this is so, and even if (as some have argued) we understand the concept of agent causation at least as well as we understand the concept of event causation, all this does nothing to diminish the mystery of free will. I will try to explain why van Inwagen thinks this by considering a particular human action. Suppose Marie wants to vote in favor of the proposal before the meeting, and that, for this reason, she raises her right hand when the chair says, "All in favor . . .?" Suppose that one of the causal antecedents of her hand's rising was a certain event in her brain that was undetermined by past events, that the state of her body and her immediate environment at the moment this brain event occurred was causally sufficient for her hand's rising, that if this event had not occurred, her hand would not have risen, and that she, Marie, a particular member of the metaphysical category 'substance' or 'continuant', was the cause—that is to say, the agent cause—of that crucial brain event. The friends of agent causation, if van Inwagen understands them, believe that these suppositions are sufficient for her having freely raised her hand. If that is so, these suppositions must entail the following proposition: at some moment shortly before Marie raised her hand, she was able to raise her hand and she was able not to raise her hand. But van Inwagen doesn't see why this entailment should be supposed to hold. In fact, he thinks he sees a good argument for the conclusion that it was not up to her whether her hand rose. Suppose God were

miraculously to return the world to precisely the state it was in, say, one minute before Marie raised her hand, and that he then allowed affairs once more to proceed, without any further miracles. What would happen? What would Marie do? Well, if her raising her hand was a free act, and if free will is incompatible with determinism, then we can't say. We can say only that she might have raised her hand and might not have raised her hand. If God were to cause this episode to be thus "replayed" a very large number of times, it might turn out that she raised her hand in thirty percent of the replays and refrained from raising it in seventy percent of the replays. This much is a simple consequence of incompatibilism, and it brings one of the main reasons philosophers become compatibilists into stark relief. It seems to lead us inescapably to the conclusion that on each particular replay, what Marie does on *that* occasion is a mere matter of chance. And if there are no replays, if there is only one occasion on which Marie is in this situation, it seems to lead us just as inescapably to the conclusion that on *that* one occasion what Marie does is a mere matter of chance. And if it is a mere matter of chance whether Marie raised her hand, then it cannot have been true beforehand that Marie was both able to raise her hand and able to refrain from raising her hand, for to have both these abilities would be to be able to determine the outcome of a process whose outcome is due to chance. It is true that we have, by stipulation, inserted into this process, this process whose outcome is due to chance, an episode of agent causation. But, if I may so express myself, so what? That doesn't change the fact that the outcome of that process was due to chance. If God caused Marie's decision to be replayed a very large number of times, sometimes (in thirty percent of the replays, let us say) Marie would have agent-caused the crucial brain event and sometimes (in seventy percent of the replays, let us say) she would not have. Surely, then, whether she agent-caused the brain event was a mere matter of chance? Whether her deliberations were followed by her agent-causing the brain event was, it would seem, a matter of chance; Marie, therefore, cannot have been both able to agent-cause the brain event and able to refrain from agent-causing the brain event, for to have both these abilities would be to be able to determine the outcome of a process whose outcome was due to chance—an impossible ability. I conclude that even if an episode of agent causation is among the causal antecedents of every voluntary human action, these episodes do nothing to undermine the *prima facie* impossibility of an undetermined free act. Postulating agent causation, therefore, does

nothing to diminish the mystery of free will. Van Inwagen's conclusion is that incompatibilists had better abandon the concept of agent causation, and seek a resolution of the mystery of free will elsewhere—if, indeed, there is an 'elsewhere'.

Notes

1. See van Inwagen (1983, 63–64).
2. See his classic essay, Frankfurt (1969).
3. Van Inwagen (1978). A somewhat condensed version of the arguments of van Inwagen (1978) can be found in van Inwagen (1983, 161–182).
4. See van Inwagen (2000).
5. At any rate, he is on record as saying things that commit him to this position. In van Inwagen (1983, 57), he said, "I am quite sure that any specific and detailed objection to one of the arguments can be fairly easily translated into specific and detailed objections to the others; and I think that any objection to one of the arguments will be a good objection to *that* argument if and only if the corresponding objections to the others are good objections to *them*." One of the arguments referred to as "the arguments" in this quotation is the argument that explicitly appeals to Rule β ; and he has said (van Inwagen 1989, 405) that if that argument is unsound, its unsoundness must be due to the invalidity of Rule β . And he has said (van Inwagen 1994, 95), "My position is that all (logically adequate) arguments for incompatibilism must make some sort of implicit or hidden or covert appeal to [Rule β]."
6. See van Inwagen (2000, 5–6).

References

- Chomsky, N. 1988. *Language and Problems of Knowledge*. Cambridge, Mass.: MIT Press.
- Davidson, D. 1973. "Freedom to Act." In T. Honderich, ed., *Essays on Freedom of Action*. London: Routledge and Kegan Paul.
- Finch, A. and T. A. Warfield. 1998. "The *Mind* Argument and Libertarianism." *Mind* 107: 516–528.
- Frankfurt, H. G. 1969. "Alternate Possibilities and Moral Responsibility." *Journal of Philosophy* 66: 828–839.
- Ginet, C. 1966. "Might We Have No Choice?" In K. Lehrer, ed., *Freedom and Determinism*. New York: Random House.

- McGinn, C. 1993. *Problems in Philosophy: The Limits of Inquiry*. Oxford: Blackwell.
- McKay, T. and D. Johnson. 1996. "A Reconsideration of an Argument against Compatibilism." *Philosophical Topics* 24: 113–122.
- Montague, R. 1974. "Deterministic Theories." In R. H. Thomason, ed., *Formal Philosophy: Selected Papers of Richard Montague*. New Haven, Conn.: Yale University Press.
- van Inwagen, P. 1974. "A Formal Approach to the Problem of Free Will and Determinism." *Theoria* 40: 9–22.
- . 1975. "The Incompatibility of Free Will and Determinism." *Philosophical Studies* 27: 185–199. Reprinted in G. Watson, ed., *Free Will*. New York: Oxford University Press, 1983; in D. Pereboom, ed., *Free Will*. Indianapolis, Ind.: Hackett, 1997; in L. Waddell Ekstrom, ed., *Agency and Responsibility: Essays on the Metaphysics of Freedom*, Boulder, Colo.: Westview Press, 2000.
- . 1978. "Ability and Responsibility." *Philosophical Review* 87: 201–224. Reprinted in John Martin Fischer, ed., *Moral Responsibility*. Ithaca, N.Y.: Cornell University Press.
- . 1983. *An Essay on Free Will*. Oxford: Clarendon Press.
- . 1985a. "On Two Arguments for Compatibilism." *Analysis* 45: 161–163.
- . 1985b. "Compatibilistic Reflections." *Australasian Journal of Philosophy* 63: 349–353.
- . 1989. "When is the Will Free?" In J. Tomberlin, ed., *Philosophical Perspectives 3: Philosophy of Mind and Action Theory*. Atascadero, Calif.: Ridgeview. Reprinted in T. O'Connor, ed., *Agents, Causes, and Events: Essays on Indeterminism and Free Will*. Oxford: Oxford University Press, 1995.
- . 1994. "When the Will Is Not Free." *Philosophical Studies* 75: 95–113.
- . 1997. "Critical Notice: John Martin Fischer's *The Metaphysics of Free Will*." *Philosophical Quarterly* 47: 373–381.
- . 1999. "Moral Responsibility, Determinism, and the Ability to Do Otherwise." *Journal of Ethics* 3: 341–350.
- . 2000. "Free Will Remains a Mystery." The Eighth *Philosophical Perspectives* Lecture. In J. Tomberlin, ed., *Philosophical Perspectives 14: Freedom and Action*. Oxford: Blackwell. Reprinted in R. Kane, ed., *The Oxford Handbook of Free Will*. Oxford: Oxford University Press, 2002.