

How to Think about the Problem of Free Will

Peter van Inwagen

Published online: 21 August 2008
© Springer Science+Business Media B.V. 2008

Abstract In this essay I present what is, I contend, the free-will problem properly thought through, or at least presented in a form in which it is possible to think about it without being constantly led astray by bad terminology and confused ideas. Bad terminology and confused ideas are not uncommon in current discussions of the problem. The worst such pieces of terminology are “libertarian free will” and “compatibilist free will.” The essay consists partly of a defense of the thesis that the use of these phrases by writers on the problem of free will can only generate conceptual confusion and partly of a formulation of the problem that does not make use of them. I contend that this formulation is neutral with respect to the historically important positions on free will (e.g., compatibilism and incompatibilism).

Keywords Compatibilism · Compatibilist free will · Determinism · Free will · Incompatibilism · Libertarian free will · Libertarianism

Perhaps we should begin with this question: What *is* the “problem of free will”? Like those other great “problem” phrases of which philosophers are so fond—“the mind-body problem,” “the problem of universals,” “the problem of evil”—this phrase has no clear referent. There are obviously a lot of philosophical problems about free will, but which of them, or which combination of them, is *the* problem of free will? I will propose an answer to this question, but this proposal can be no more than that, a proposal. I propose that we understand the problem of free will to be the following problem.

There are seemingly unanswerable arguments that (if they are indeed unanswerable) demonstrate that free will is incompatible with determinism.

P. van Inwagen (✉)
Department of Philosophy, University of Notre Dame, Notre Dame, IN 46556, USA
e-mail: Peter.vaninwagen.1@nd.edu

And there are seemingly unanswerable arguments that (if indeed ...) demonstrate that free will is incompatible with indeterminism. But if free will is incompatible both with determinism and indeterminism, the *concept* “free will” is incoherent, and the *thing* free will does not exist. There are, moreover, seemingly unanswerable arguments that, if they are correct, demonstrate that the existence of moral responsibility entails the existence of free will, and, therefore, if free will does not exist, moral responsibility does not exist either. It is, however, evident that moral responsibility does exist: if there were no such thing as moral responsibility nothing would be anyone’s fault, and it is evident that there are states of affairs to which one can point and say, correctly, to certain people: “That is *your* fault.” It must, therefore, be that at least one of the following three things is true:

- The seemingly unanswerable arguments for the incompatibility of free will and determinism are in fact answerable; these arguments are fallacious;
- The seemingly unanswerable arguments for the incompatibility of free will and indeterminism are in fact answerable; these arguments are fallacious;
- The seemingly unanswerable arguments for the conclusion that the existence of moral responsibility entails the existence of free will are in fact answerable; these arguments are fallacious.

The “problem of free will” is just this problem (this is my proposal): to find out which of these arguments is fallacious, and to enable us to identify the fallacy or fallacies on which they depend.

Having set out a philosophical problem, and having, tentatively, identified this problem with the problem of free will, I will define the important terms that occur in the statement of the problem.

Since I am presenting a paper that consists largely of advice, I will preface these definitions with some advice about how to frame definitions.

(1) Define every term you use that is neither a word or phrase of ordinary language nor a technical term of some discipline other than philosophy (in which discipline, one supposes, it will have been given an adequate definition). Do not waive this requirement in the case of some term simply because philosophers use that term a lot. If you think that some term you will use has been given an adequate definition in the philosophical literature, repeat that definition.

This, at any rate, is the ideal. It will often be impossible to write an essay (as opposed to a very long book) that conforms to this ideal. But, insofar as one’s essay does not conform to this ideal, one has issued a promissory note: definitions of all one’s technical terms should be available on request.

(2) I have mentioned defining terms, but that was loose language. Here is a second piece of advice about framing definitions: define sentences, not terms. Do not, for example, define “cause” or “causation” or “causality;” rather, define “*x* is the cause of *y*” or “*x* is a cause of *y*” or “*x* causes *y*.” Do not define “knowledge;” rather, define “*x* knows that *p*.” And a definition of, e.g., “*x* causes *y*” should take this form: a sentence that can replace “*x* causes *y*” at all its occurrences, a sentence

in which “ x ” and “ y ” and no other variables are free, together with a specification of the kinds of object over which “ x ” and “ y ” range. A definition of “ x knows that p ” should contain the free variable “ x ” and the schematic sentence-letter “ p ” and no other free variables or schematic letters.

This stern requirement—the “Chisholm requirement” so to call it—can be softened in one way: it is permissible to define nouns and noun-phrases (“terms” in the strict sense of the word) if they are the names of theses or propositions. Thus, a proper definition can consist of the phrase, “Ethical naturalism is the thesis that” followed by a declarative sentence, or a series of them, that spells out what the philosopher offering the definition takes to be the content of the thesis called “ethical naturalism.” This is a softening of the Roderick Chisholm requirement, and not a contradiction of it, because, like the pristine Chisholm requirement, it demands a definition whose *definiens* contains a complete declarative sentence (or a series of them). Why, you may ask, do I “privilege” *declarative sentence* over the many other syntactic categories in this way? My answer is this: the declarative sentence is the natural unit of clear statement.

The following series of definitions conforms to my second requirement. Note, for example, that my first definition is a definition of “the free-will thesis” and not of “free will”¹ (Unfortunately, one of the definitions does not conform to my *first* requirement: the definition of “determinism” contains the phrases “the laws of nature” and “determine a unique future,” neither of which is either ordinary language, a technical term of some science like physics or geology, or defined. My excuse is the one I have mentioned: this is a paper, not a book. I remind you, however, that I *have* written a book).

These definitions serve to explain the system of concepts everyone who thinks about the free-will problem should use—or so I say.

- The free-will thesis is the thesis that we are sometimes in the following position with respect to a contemplated future act: we simultaneously have both the following abilities: the ability to perform that act and the ability to refrain from performing that act (This entails that we *have been* in the following position: for something we did do, we were at some point prior to our doing it able to refrain from doing it, able not to do it).²

¹ The phrase “free will”—whether it occurs by itself or within the phrase “the free-will thesis”—hardly exists except as a philosophical term of art. Its non-philosophical uses are pretty much confined to the phrase “of his/her own free will” which means “uncoerced.” When, in the movie *Devil’s Advocate*, Keanu Reeves says to Al Pacino, “But suppose I sell my soul to you, and then repent on my deathbed ...” and Pacino, the Devil, replies, “Yeah—free will. That one’s a *bitch*,” the latter is using a term from philosophy: a technical term from philosophy that, by way of theology, has achieved everyday currency with little if any distortion of the meaning it has in philosophy, a very uncommon fate for a technical philosophical term. “Free will,” although a technical term, was, for centuries, bandied about without any real attempt at definition—like “continuous function” in pre-nineteenth-century mathematics. It is my position that the definition of “the free-will thesis” that follows in the text provides a philosophically adequate definition of this old technical term.

² Whatever you do, do not define “free will” this way: “Free will is whatever sort of freedom is required for moral responsibility” (or “Free will consists in having whatever sort of access to alternative possibilities is required for moral responsibility”).

- Determinism is the thesis that the past and the laws of nature together determine, at every moment, a unique future (The denial of determinism is indeterminism).
- Compatibilism is the thesis that determinism and the free-will thesis could both be true (And incompatibilism is the denial of compatibilism).³
- Libertarianism is the conjunction of the free-will thesis and incompatibilism (Libertarianism thus entails indeterminism).
- Hard determinism is the conjunction of determinism and incompatibilism (Hard determinism thus entails the denial of the free-will thesis).
- Soft determinism is the conjunction of determinism and the free-will thesis (Soft determinism thus entails compatibilism).

I must emphasize that I intend these definitions to be entirely neutral as regards competing positions about the relations between free will and determinism and entirely neutral as regards competing accounts of the nature of free will. In principle, of course, as a matter of logical *theory*, a definition cannot favor one thesis over its logical contraries. A definition is a declaration about how words are to be used, and a rose would smell as sweet even if we all agreed to use the word “rose” to mean “chamber pot.” Roses and their scents are elements of reality, and reality is serenely indifferent to the ways in which human beings use words. In practice, however, one has to admit that—owing to the perverse human tendency to confuse words and things—definitions can be tendentious. I contend that my definitions are not tendentious. I contend that agreement to abide by the above set of definitions will not confer any hidden advantage on any position that figures in philosophical debates about free-will—including, of course, my own. In defense of this contention, I cite the fact that the finest essay that has ever been written in defense of compatibilism—possibly the finest essay that has ever been written about any aspect of the free-will problem—, David Lewis’s “Are We Free to Break the Laws?”⁴ opens with definitions of “determinism,” “soft determinism,” “hard determinism,” and “compatibilism” that are equivalent to the definitions I have set out above (Lewis gave no definition of “libertarianism” because he had no use for the term. I have no use for it either, but, in point of fact, my series of definitions contains definitions of several terms that I have no use for. But some people do use these terms, and I am laying down the law about what someone who does use them should mean by them).

Since “moral responsibility” figures prominently in my statement of the free-will problem, one might expect that at this point I should define this term, or at least define some sentence or sentences in which it occurs—“*x* is morally responsible for *y*,” perhaps. I will not do this. If I *did* offer a definition in this general area, it would be something like this:

x is morally responsible for the fact that *p* =_{df} It is *x*’s fault that *p*.

But so much confusion attends the phrase “moral responsibility” (the confusion is of our own making; as George Berkeley said, “... we have first raised a dust, and

³ Whatever you do, do not use “compatibilism” as a name for the thesis that *moral responsibility* is compatible with determinism. This can only cause confusion. If you must have a name for this thesis, invent a new one: “MR-compatibilism” or some such.

⁴ Lewis (1981, pp. 113–121).

then complain we cannot see”) that I despair of straightening it all out in a paper that is not devoted to that topic alone. In the sequel, instead of discussing free will and moral responsibility, I will discuss a simpler (but obviously closely related) topic, free will and negative moral judgments. A fully adequate discussion of the problem of free will, however, would require a discussion of the relations that hold between free will and moral responsibility.

The last three “-isms” in my list, libertarianism, hard determinism, and soft determinism, are conjunctions of more fundamental theses. Although the terms “libertarianism,” “hard determinism,” and soft “determinism” are perfectly well defined, I very strongly recommend that philosophers never use them—except, of course, when they are forced to because they are discussing the work of philosophers who have been imprudent enough to use them (This was what I was alluding to when I said that I had no use for several of the terms for which I provided definitions). Writers on free will who do not take my advice on this matter are continually saying things that they would be better off not saying—and they would not say these unfortunate things, they would automatically avoid saying them, if they confined their list of technical terms to “the free-will thesis,” “determinism,” “compatibilism” and “incompatibilism.” A minor example is this: There is a tendency among writers on free will to oppose “compatibilism” and “libertarianism;” but the fundamental opposition is between compatibilism and incompatibilism. Here is a major example (not entirely unconnected with my minor example). Philosophers who use the term “libertarianism” apparently face an almost irresistible temptation to speak of “libertarian free will.”

What is this libertarian free will they speak of? What does the phrase “libertarian free will” mean? Since I am discussing the usage of others, of certain people who have perversely chosen not to follow my advice about the proper way to frame definitions, I am going to have to try to answer a question about the meaning of an abstract noun-phrase. And this faces me with some very real semantical difficulties. I generally have trouble understanding writers who do not take my advice about how to frame definitions. Noun-phrases like “free will” and “compatibilist free will” and “libertarian free will” are particularly difficult for me. I find it difficult to see what sort of thing such phrases are supposed to denote. In serious philosophy, I try never to use an abstract noun or noun-phrase unless it is clear what ontological category the thing it purports to denote belongs to. For many abstract noun-phrases, it is not at all clear what sort of thing they are supposed to denote, and I therefore try to use such phrases only in introductory passages, passages in which the reader’s attention is being engaged and a little mush does not matter. In serious philosophical argument about free will, I try to restrict my use of abstract noun-phrases to phrases that indisputably denote *propositions* (such as “the free-will thesis” and “determinism”). If I had to guess what sort of thing “free will” etc. denoted, one guess I might make is that they denote a certain *properties*. On that guess, “free will” is a name for the property *is on some occasions able to do otherwise*. But then what properties are denoted by “compatibilist free will” and “libertarian free will”? Well, presumably, e.g., “libertarian free will” denotes whatever it is that “free will” denotes when libertarians use it. And “free will” denotes the same thing when

anyone uses it. So, if “free will” denotes the property *is on some occasions able to do otherwise*, there is really nothing for “libertarian free will” to denote but that same property. Another guess about what “free will” might denote is: it denotes a certain power or ability: the power or ability sometimes to do otherwise than what one in fact does (That is, of course, *another* guess only if a power is something other than a property). If that is what “free will” denotes, a power, then, by an argument parallel to the above, “libertarian free will” and “compatibilist free will” denote that same power.

The operative word in both guesses is “able”—as in “Jill says she is able to do what I have asked.” Many philosophers, in attempting to spell out the concept of free will, use the phrase “could have done otherwise.” I did so myself in *An Essay on Free Will*.⁵ Nowadays, however, I very deliberately avoid this phrase. I avoid it because “could have done otherwise” is ambiguous and (experience has shown) its ambiguity has caused much confusion in discussions of free will. My advice to all philosophers who write about free will in English is that they should avoid this phrase as well. A whole chapter of Daniel Dennett’s first book on free will⁶ was written to no purpose because he did not realize that “could have done” sometimes means “might have done” (and this “might” is itself ambiguous: it has both an ontological and an epistemic sense) and sometimes “was able to do.” Austin was very clear about this in his classic paper “Ifs and Cans,”⁷ but subsequent writers on free will have not learned what he had to teach⁸ (I cannot resist the temptation to mention a passage in Austin’s essay that delights me because it is such a perfect reflection of a vanished culture. Being an Englishman with a pre-war public-school education speaking to an audience of his cultural peers, Austin found it natural to explain the two senses of “could have” by contrasting a case in which a Latin speaker would have said “*potui*” and a case in which he would have said “*potuissem*.”⁹). This ambiguity in the phrase “could have done otherwise” has led a considerable body of philosophers to think that to say that someone could have done otherwise is to imply something having to do with “alternative possibilities,” to imply that the person’s act was undetermined. And, indeed, when it means “might have done otherwise” (in the ontological, as opposed to the epistemic, sense of “might have”), that is just what “could have done otherwise” *does* imply. But those who have defined free will in terms of the phrase “could have done otherwise” were using the phrase in its other sense: “was able to do otherwise.” They would have done better simply to have avoided the ambiguous phrase and to have used “was able to.”

⁵ van Inwagen (1983).

⁶ Dennett (1984, Ch. 6).

⁷ Austin (1961, pp. 153–180).

⁸ Dennett (1984, Ch. 6), cited in footnote 6, contains a rather lengthy discussion of “Ifs and Cans.” In this discussion Dennett presents telling criticisms of some of Austin’s arguments (I would call your attention to his excellent discussion of the notorious “I could have holed it” example). But Dennett does not see the point I am calling attention to—although he sometimes dances maddeningly close to it. For more on this topic, see van Inwagen (1988).

⁹ Austin (1961, p. 163).

All the phrases that have been used in definitions of “free will” (and in statements of the free-will thesis) can be defined in terms of, or dispensed with in favor of, “able.” For example, the much-used phrase “within one’s power” can be defined like this: “It is within x ’s power to” means “ x is able to.” Having said this about the word “able” I want to make what seems to me to be an important point, a point that is, in fact, of central importance if one wishes to think clearly about the freedom of the will: compatibilists and incompatibilists mean the same thing by “able.” And what do both compatibilists and incompatibilists mean by “able”? Just this: what it means in English, what the word means. And, therefore, “free will,” “incompatibilist free will,” “compatibilist free will” and “libertarian free will” are four names for one and the same thing. If this thing is a property, they are four names for the property *is on some occasions able to do otherwise*. If this thing is a power or ability, they are four names for the power or ability to do otherwise than one in fact does.

All the compatibilists I know of believe in free will. Many incompatibilists (just exactly the libertarians: that is how “libertarian” is defined) believe in free will. And it is one and the same thing they believe in. Compatibilists say that the existence of this thing (whose conceptual identity is determined by the meaning of the English word “able,” or of some more-or-less-equivalent word or phrase in some other language) is compatible with determinism; incompatibilists say that the existence of this thing is incompatible with determinism. If Alice used to be an incompatibilist and has been converted by some philosophical argument to compatibilism, she should describe her intellectual history this way: “I used to think that free will was incompatible with determinism. I was blind but now I see: Now I see that *it* is compatible with determinism.” And her use of “it” does not have to be apologized for: this very thing she used to think was incompatible with determinism, she now thinks is compatible with determinism (Compare: I used to think that knowledge was incompatible with the logical possibility of a Universal Deceiver. Now I see that *it* is compatible with the logical possibility of such a being). What Alice should *not* say is this:

I used to think that free will was one thing, a thing incompatible with determinism. Now I think it is another thing—I mean I think that the words “free will” are a name for another thing—, a thing compatible with determinism. The thing I used, incorrectly, to call “free will” *is* incompatible with determinism; I was right to think it was incompatible with determinism. But it does not exist (I mean no agent has it), and it could not exist, and if it did exist, it would not be right to call it “free will.”

Talk of “libertarian free will” is therefore at best useless. Taking this phrase seriously as a denoting phrase would be like taking “materialist pain” (or “pain according to the materialists”) seriously as a denoting phrase. Suppose someone said that “pain according to the materialists” purported to denote some particular sort of brain process, and that it therefore did not exist unless materialism was the right philosophy of mind. That would be silly. Materialists and dualists and idealists all use “pain” to refer to the same thing.

Let me break off at this point and remark that the fallacy of which I am accusing those who speak of libertarian free will is a very general sort of fallacy whose

pernicious effects extend far beyond the problem of free will. The fallacy need not involve abstract nouns like “free will.” I have seen a very pure instance of it that involves the most concrete of all nouns, “God.” I have seen a letter to the editors of the *Chicago Tribune* the author of which contended that it could not be that (as had been maintained by the author of an earlier letter) Christians and Muslims worshiped the same God, since the Christian God was incarnate in Jesus Christ and the Muslim God was not. But, of course, “the Christian God” can only mean “the God worshiped by Christians” and “the Muslim God” can only mean “the God worshiped by Muslims.” And the three phrases “God” and “the God worshiped by Christians” and “the God worshiped by Muslims” all denote the same being. Even if there is no God, this statement is true, as it were, counterfactually: if any of the three phrases *did* denote something, the other two would also denote it (As to the letter-writer’s logic, one might as well argue that since, in the writings of Malcolm Muggeridge, Mother Theresa was a saint, and, in the writings of Christopher Hitchens, Mother Theresa was a charlatan, Muggeridge and Hitchens were therefore not writing about the same woman).

If the materialists (to return to our primary example) are wrong about pain being a physical process, they have nevertheless been referring to it all along. On conversion to some other view, they should not say, “I see now that there is no such thing as what I called ‘pain’.” They should say, “I see now that pain does not have some of the properties I thought it had; for one thing it is not a physical process.” And libertarians who become compatibilists should not say, “I see now that there is no such thing as what I called ‘free will’.” They should say, “I see now that free will does not have some of the properties I thought it had; for one thing, it is not incompatible with determinism.”

Use of the phrase “libertarian free will” can lead critics of libertarianism into confusion about what it is that libertarians believe. I can only tear my hair when I read the following statement¹⁰:

van Inwagen has said of an argument concluding that libertarian free will is incompatible with *indeterminism*, “I fervently hope that there is something wrong with it.”

I would never (except in discussions of the work of others) use the phrase “libertarian free will;” and nothing I have written can be accurately paraphrased using that phrase. The conclusion of the argument to which Professor Baker refers¹¹ was that free will *simpliciter*, that one free will that both compatibilists and incompatibilists believe in, was incompatible with indeterminism. The argument that I said I hoped was defective was a version of the “*Mind Argument*.” I will discuss the *Mind Argument* later in this essay. Here I will note only that that most of the proponents of the *Mind Argument* have been *compatibilists*, who have used it to defend the conclusion that free will—not compatibilist free will, not libertarian free will, but free will full stop—is incompatible with indeterminism (The *Mind Argument* is perhaps the sharpest arrow in the compatibilist’s quiver. It is a sharp

¹⁰ Baker (2003, p. 476).

¹¹ See van Inwagen (2000). The words Baker quotes appear on p. 175 of Kane (2002).

arrow for two reasons: (i) it is very hard to answer, and (ii) it follows from its conclusion that if free will is possible, free will is compatible with determinism).

The pernicious phrase “libertarian free will,” has led critics of libertarianism to misunderstand not only what libertarians believe but the motives that underlie their beliefs. Professor Baker, for example, has supposed (in the article I have cited) that libertarians “want” libertarian free will—which means, presumably, both that they want to have it (as they no doubt want to be wise and virtuous) and that, in some more purely intellectual and disinterested way, they want libertarianism to be the correct theory of free will.¹² These are, I suppose, natural things for someone to suppose if that person thinks that there is something called “libertarian free will” and that it is a different thing from “compatibilist free will.” But, as I have said, insofar as it makes sense to treat “libertarian free will” as a denoting term, it has to be regarded as a name (a rather misleading name) for free will—for free will *simpliciter*, free will *tout court*, free will full stop, free will period. And this one thing, free will, is what *both* libertarians and soft determinists want to have. It is simply not true that there are two distinct things, libertarian free will and compatibilist free will, and that libertarians want the one and do not regard the other as worth having (regard what the compatibilists offer as “free will” as, in Immanuel Kant’s words, a wretched subterfuge). There is, I concede, some historical justification for this. Kant, if I read him right, saw David Hume as offering a substitute for free will, an *ersatz* free will that was a mere pretender to the name. But this is not the way present-day incompatibilists view what their compatibilist colleagues call free will. We present-day incompatibilists see the free will that compatibilists believe in as the genuine article; their only mistake, in our view, is to suppose that it, the genuine article, is compatible with determinism.

Let us turn from what libertarians want to have to what they want to be true. Do libertarians want libertarianism to be true? Well, libertarianism is the conjunction of the free-will thesis and incompatibilism. To want libertarianism to be true, therefore, would be to want both the free-will thesis and incompatibilism to be true. I will stipulate, as the lawyers say, that libertarians want the free-will thesis to be true. (And who would not? Even hard determinists, or most of them, seem to regard the fact—they think it is a fact—that we do not have free will as a matter for regret.) But do libertarians want incompatibilism to be true? Perhaps some do. I can say only that *I* do not want incompatibilism to be true. Just as hard determinists regard

¹² Baker’s essay is primarily an argument for the conclusion that *Christian* philosophers should not be libertarians. In the matter of motives and desires, she is primarily concerned to defend the position that Christian philosophers should not want to have “libertarian free will” or want libertarianism to be true. All but one of the reasons for wanting libertarianism to be true that she canvases and rejects in the section of the paper called “Motives for Belief in Libertarian Free Will,” however, have nothing in particular to do with theism in general or Christianity in particular: Libertarian free will is required for moral responsibility; without libertarian free will, we would be puppets; libertarian free will is presupposed by the principle “ought” implies ‘can’.” The single “motive for belief in libertarian free will” she considers that would be a motive only for Christians (or at any rate, only for theists) is that libertarian free will is (supposedly) essential to an effective reply to the argument from evil. In this essay, I am concerned to address only the motives of libertarians that are essentially connected with what I am calling “the problem of free will.” I am not concerned to address motives that a libertarian might have that arose from beliefs he or she had that (like theism) were not consequences of the premises that generate the problem of free will.

the non-existence of free will as a matter for regret, I regard the fact—I think it is a fact—that free will is incompatible with determinism as a matter for regret. But reason has convinced me that free will is incompatible with determinism, and I have to accept the deliverances of reason, however unpalatable they may be. I should think that any philosopher in his or her right mind would *want* compatibilism to be true. It would make everything so *simple*. But we cannot always have what we want and things are not always simple.

I have used Professor Baker as an example of a philosopher who has misunderstood what libertarians believe and want because she provides a recent and very clear example of such a philosopher. But heaven forbid that I should be thought to have implied that she is unique or even unusual in this respect. I could cite similar mistakes on the part of many others. Consider Dennett, a philosopher always worth considering. The same sort of mistake is on display in his chapter of *Brainstorms*, “On Giving Libertarians What They Say They Want” (and in the essay whose title it is¹³). I repeat: What libertarians want is identical with what soft determinists want: free will, the ability to do otherwise (There *is* something we libertarians want that soft determinists do not want. We libertarians want to know what is wrong with the well-known arguments for the incompatibility of free will and indeterminism. Soft determinists do not want this, or most of them do not, because they, or most of them, do not think that there is anything wrong with those arguments. But this is not the thing that libertarians say they want that Dennett was talking about). We may consider also in this connection the subtitle of Dennett’s *Elbow Room*, “*The Varieties of Free Will Worth Wanting*.” There is only one variety of free will worth wanting because there is only one variety of free will: the ability to do otherwise. And everyone wants that, both those who think human beings have it (libertarians and soft determinists and compatibilists who are not determinists) and most of those who think they do not (most hard determinists).

I should at some point, and this is as good a point as any, make a remark about the phrase “libertarian free will” that I have found it rather hard to find a natural place for in my discussion of the phrase. There is a way to define “libertarian free will” that does not have the consequence that “libertarian free will” is simply another name for free will *simpliciter*. One *might* define “libertarian free will” like this:

x has libertarian free will =_{df} x has free will *simpliciter* and free will *simpliciter* is incompatible with determinism.

Similarly, one might offer this definition of compatibilist free will:

x has compatibilist free will =_{df} x has free will *simpliciter* and free will *simpliciter* is compatible with determinism.

Given these definitions, libertarian free will is not identical with compatibilist free will. Libertarian free will and compatibilist free will are logical contraries: they cannot both exist.

¹³ Dennett (1978, pp. 286–299).

I can only say that no libertarian has ever used “free will” to mean “libertarian free will” in this sense, and that no compatibilist has ever used “free will” to mean “compatibilist free will” in this sense. It is therefore hard to see what point these definitions might have. “Libertarian free will” in the present sense is by definition incompatible with determinism, and “compatibilist free will” is by definition compatible with determinism. If, therefore, libertarians had used “free will” to mean “libertarian free will” it is hard to see why they would have bothered to offer arguments for the incompatibility of free will and determinism. And if compatibilists had used “free will” to mean “compatibilist free will” it is hard to see why they would have bothered to offer arguments for the compatibility of free will and determinism. (Insofar as they have. I concede that my argument on this point is somewhat weakened by the fact that it is pretty rare for a compatibilist actually to present an argument for compatibilism. The more usual procedure among compatibilists is to treat compatibilism as the “default” position concerning the relation between free will and determinism, and—if they do even this much—to try to refute the standard arguments for incompatibilism.) I would also point out that, if the words are given this sense, I do not want “libertarian free will;” I want free will *simpliciter*, of course, but I should much rather have compatibilist free will than libertarian free will. It would make everything so simple. Unfortunately, as I see matters, I have free will *simpliciter* if and only if I have libertarian free will (in the present bizarre sense of “libertarian free will”). That is to say, as I see matters, free will *simpliciter* is incompatible with determinism.

What I have presented, I contend, is the free-will problem properly thought through—or at least presented in a form in which it is possible to think about it without being constantly led astray by bad terminology and confused ideas. Let me now restate the problem of free will, this time using only “basic” terminology, the undefined terms that occur in my definitions.

The following two theses are *prima facie* incompatible:

- (1) We are sometimes in the following position with respect to a contemplated future act: we simultaneously have both the following abilities: the ability to perform that act and the ability to refrain from performing that act.
- (2) The past and the laws of nature together determine, at every moment, a unique future.

These two theses are *prima facie* incompatible because the premises of the Consequence Argument (an argument whose logical validity no one disputes) are *prima facie* true. “The Consequence Argument” is my name for the standard argument (various more-or-less equivalent versions of the argument have been formulated by C.D. Broad, Chisholm, David Wiggins, Carl Ginet, James Lamb, and myself) for the incompatibility of (1) and (2). It is beyond the scope of this paper seriously to discuss the Consequence Argument. I will, however, make a sociological point. Before the Consequence Argument was well known (Broad had formulated an excellent version of it in the 1930s, but no one was listening¹⁴), almost all

¹⁴ Broad (1952, pp. 195–217). The informational content of the cited chapter constituted Broad’s Cambridge Inaugural Lecture as Knightsbridge Professor in 1934.

philosophers who had a view on the matter were compatibilists. It is probably still true that most philosophers are compatibilists. But it is also true that the majority of philosophers who have a specialist's knowledge of the ins and outs of the free-will problem are incompatibilists.¹⁵ And this change is due entirely to the power, the power to *convince*, the power to move the intellect, of the Consequence Argument. If, therefore, the Consequence Argument is fallacious (in some loose sense; it certainly contains no *logical* fallacy), the fallacy it embodies is no trivial one. Before the Consequence Argument was well known, most philosophers thought that incompatibilists (such incompatibilists as there were) were the victims of a logical "howler" that could be exposed in a paragraph or two. No one supposes that now (I mean this to be true in the sense that "No one now believes in the Divine Right of Kings" is true. *Some* people, of course, do believe in the Divine Right of Kings). The *prima facie* incompatibility of (1) and (2) is the first of three components of the problem of free will. Now the second.

The following proposition is *prima facie* true.

- (3) Necessarily: If one is contemplating some possible future act, and if the past and the laws of nature do *not* together determine that one shall perform that act, then one is unable to perform that act.

Proposition (3) is *prima facie* true because the *Mind* Argument, whose conclusion it is, is, *prima facie*, a cogent argument. "The *Mind* Argument" is my name for an argument that may be loosely stated like this: If what one does not follow deterministically from one's previous states, then it is the result of an indeterministic process, and (necessarily) one is unable to determine the outcome of an indeterministic process.¹⁶

Now consider this fact. Propositions (1) and (3) jointly entail proposition (2). If, therefore, (3) is true, and if (1) entails the denial of (2), then (1) is a necessary falsehood. If (1) is incompatible with (2), therefore, and if (3) is true, it is impossible for there to be agents who are able to do anything other than what they in fact do (Indeed—whether this is so depends on how some tricky questions are answered—(3) may imply that, if the world is indeterministic, agents are not even able to do the things that they in fact *do*).

We have, therefore, a *prima facie* case for the impossibility, the ground-floor or metaphysical impossibility, of the following proposition: When we deliberate about which of two actions, *A* and *B*, to perform, we are sometimes able to perform *A* and able to perform *B*. We have, to speak loosely, a *prima facie* case for the metaphysical impossibility of free will. Broad thought that the reasoning I have set out (or something very much like it) was not only *prima facie* correct but correct *tout court*, and that this proposition is in fact impossible. Might we not simply agree

¹⁵ I.e., incompatibilists in the proper sense of the term (see footnote 3, above). There are, of course, philosophers who believe that *free will* and determinism are incompatible and who also believe that *moral responsibility* and determinism are compatible.

¹⁶ The name "The *Mind* Argument" is due to the fact that between 1930 and 1960, versions of the argument appeared regularly in that august philosophical journal. One example is Hobart (1934, pp. 1–27). I have myself presented a version of the *Mind* Argument in van Inwagen (2000).

with him? Might we not simply agree that free will (not “libertarian free will” but free will *simpliciter*) is metaphysically impossible?

This question brings us to the third component of the free-will problem: an argument for the conclusion that moral responsibility requires the ability to act otherwise than we do. As I have said, however, owing to widespread confusions that attend the concept of moral responsibility, confusions I cannot address within the scope of this essay, I shall not present an argument for this conclusion. I shall instead present an argument for the conclusion that the truth of negative moral judgments about a person’s acts implies that that person was able to do something other than the act that is the object of the moral judgment (If you tell me that, in your view, moral judgments do not have truth-values, I will reply that my argument does not really require the premise that moral judgments have truth-values; the argument could be reconstructed in other terms). The strategy on display in this argument, I contend, could, with minor modifications, be used to construct an equally plausible argument for the conclusion that moral responsibility implies the ability to act otherwise.

Suppose that your friend Alice has told a lie, and that you say to her (stern moralist that you are),

(a) You ought not to have lied.

Making statement (a), it would seem, commits you to the truth of

(b) You ought either to have told the truth or to have remained silent.

And (b), in its turn, commits you to the truth of

(c) You were able either to tell the truth or [inclusive] remain silent.

Note in connection with statement (c) that we always accept “I was not able to do x ” as an excuse for not doing x —provided, of course, that we *believe* that statement. (So I say. Some might want to dispute this in certain cases in which the inability that the speaker claims is a consequence of the speaker’s own prior acts. Suppose, for example, that our Alice had replied to statement (a) by saying, “I could not help lying. I had been drinking and I always turn into a pathological liar when I have had a few. When I am drunk, I simply lose the ability to speak truthfully.” You will probably want to tell me that this fails as an excuse, and I will agree. At least I will agree to this extent: Alice cannot expect you, the stern moralist of the example, to respond by saying, “Oh, I did not know that. That is all right, then.” But I do not think that what I have conceded implies that you, the stern moralist, can properly respond by saying, “Granted, but you still ought not to have lied.” What you, in your stern moralist role, should say is rather, “Well, if drinking affects you that way, you should not drink. You do not claim that you were not able to refrain from drinking, do you?” If Alice replies that she is literally unable to resist the temptation to drink, you can—if you believe this statement—tell her that in that case she ought to avoid situations in which she might be tempted to drink. If she replies that she tries, and tries very hard and very intelligently, to avoid situations in which drink might be offered her, but contends that she was exposed to temptation in present instance because of humanly unforeseeable circumstances—well, perhaps at this point, assuming that everything Alice says is to be believed, even the sternest

moralist ought to leave off making moral judgments.”) Statement (c) commits you to the truth of

(d) You were able to do something you did not do,

and (d) commits you to the truth of (1), to the free-will thesis.

Therefore, if the free-will thesis is false, negative moral judgments are always false. (Or if moral judgments lack truth-values, negative moral judgments are always in some way out of place or inappropriate. Even if moral judgments lack truth values, there is obviously *something* wrong with telling King Canute that he ought to have succeeded in halting the advance of the tide. If agents are never able to do anything other than what they in fact do, all moral judgments share whatever defect it is that that one so prominently displays.) It is, however, undeniable that people do not always behave as they ought. (In the words of *The Book of Common Prayer*, “We have left undone those things which we ought to have done, and we have done those things which we ought not to have done.”) Denying that agents are ever able to do otherwise is therefore simply not an option.

These are the three arguments that create the problem of free will. Each of them is *prima facie* correct, but at least one of them must contain some error. But which? And where does the error (or where do the errors) lie? That is the problem of free will. I myself think that the error must lie in the *Mind* Argument. But I have not the faintest idea what the nature of the error is (Most of my fellow libertarians think that that the error in the *Mind* Argument—they agree with my conviction that that is where the error is to be found—can be exposed by reflection of the concept of “agent causation.” I cannot agree. In my view, even if agent causation exists and underlies all our free actions, this does not point to any defect in the *Mind* Argument.¹⁷ The advice I have offered in this paper has mostly been directed at compatibilists. Here is a piece of advice for incompatibilists: do not underestimate the power of the *Mind* Argument).

The problem of free will, I believe, confronts us philosophers with a great mystery. Under it our genius is rebuked. But confronting a mystery is no excuse for being in a muddle. In accusing others of muddle, I do not mean to imply that they are muddled because they do not believe what I do about free will. I do not mean to imply that they are muddled because they are compatibilists. I am an incompatibilist and Lewis was a compatibilist. But the two of us have framed the problem of free will in the same terms. I, naturally enough, do not think I am muddled, and I do not think Lewis was either. No indeed: he saw the problem with his usual crystalline clarity. Here is my closing piece of advice for compatibilists. Study “Are We Free to Break the Laws?” carefully. *That* is the way to be a compatibilist.

References

Austin, J. 1961. Ifs and cans. In *Philosophical papers*, ed. J. Urmson and G. Warnock, 153–180. Oxford: Clarendon Press.

¹⁷ My reasons for thinking this can be found in van Inwagen (2000).

- Baker, L. 2003. Why Christians should not be libertarians: An Augustinian challenge. *Faith and Philosophy* 20: 460–478.
- Broad, C.D. 1952. *Ethics and the history of philosophy*. London: Routledge & Kegan Paul.
- Dennett, D. 1978. *Brainstorms*. Cambridge: The MIT Press.
- Dennett, D. 1984. *Elbow room: The varieties of free will worth wanting*. Cambridge: The MIT Press.
- Hobart, R. 1934. Free-will as involving determination and inconceivable without it. *Mind* 43: 1–27.
- Kane, R. (ed.). 2002. *The Oxford handbook of free will*. Oxford: Oxford University Press.
- Lewis, D. 1981. Are we free to break the laws? *Theoria* 47: 113–121.
- van Inwagen, P. 1983. *An Essay on free will*. Oxford: Clarendon Press.
- van Inwagen, P. 1988. Review of Daniel Dennett, *Elbow room*. *Nouûs* 22: 609–618.
- van Inwagen, P. 2000. Free will remains a mystery. *Philosophical Perspectives* 14: 1–19.