

# Arguments for Incompatibilism

Kadri Vihvelin

*First published Tue Oct 14, 2003; substantive revision Tue Nov 13, 2007*

Determinism is a claim about the laws of nature: very roughly, it is the claim that *everything* that happens is determined by antecedent conditions together with the natural laws. Incompatibilism is a philosophical thesis about the relevance of determinism to free will: that the truth of determinism rules out the existence of free will. The incompatibilist believes that if determinism turned out to be true, it would also be true that we don't have, and have never had, free will. The compatibilist denies that determinism has the consequences the incompatibilist thinks it has. According to the compatibilist, the truth of determinism does not preclude the existence of free will. (Even if we learned tomorrow that determinism is true, it might still be true that we have free will.) The philosophical problem of free will and determinism is the problem of understanding, how, if at all, the truth of determinism might be compatible with the truth of our belief that we have free will. That is, it's the problem of deciding who is right: the compatibilist or the incompatibilist.

Why an encyclopedia entry on arguments for incompatibilism? (Why not an entry on the problem of free will and determinism?) Perhaps for this reason: until fairly recently, compatibilism was the received view and it was widely believed that arguments for incompatibilism rest on a modal fallacy or fairly obvious mistake (e.g., the mistake of confusing causation with compulsion, or the mistake of confusing descriptive with prescriptive laws) (Ayer 1954, Dennett 1984). Compatibilists have also tended to dismiss incompatibilism because of its guilt by association with a metaphysical worldview that P. F. Strawson famously dismissed as “obscure and panicky” — dualism, agent-causation (Strawson 1962). Indeed, thanks to Strawson's seminal paper, many compatibilists are convinced that the free will/determinism problem is not a *metaphysical* problem at all (as opposed to a problem about moral responsibility which arises within normative ethics or metaethics) (Wallace 1994). And even those compatibilists who regard the problem as a metaphysical problem have, for the most part, been pre-occupied with defending free will against those who argue that free will is either *impossible* or *empirically implausible* regardless of whether determinism is true or false (Wolf 1990). And so the literature on the problem of free will and determinism has come to be dominated by incompatibilists.

- [1. Definitions and Distinctions](#)
- [2. Classification of Arguments for Incompatibilism](#)
- [3. Arguments based on Intuition](#)
  - [3.1 Garden of Forking Paths Argument](#)
  - [3.2 Manipulation and Design Arguments](#)
- [4. Self-Determination and the Causal Chain argument](#)
- [5. Choice and the Consequence Argument](#)
- [Bibliography](#)
- [Other Internet Resources](#)

- [Related Entries](#)
- 

## 1. Definitions and Distinctions

In the literature, “determinism” is sometimes used as an umbrella term for a variety of different claims which have traditionally been regarded as threats to free will. Given this usage, the thesis that I am calling “determinism” (nomological determinism) is just one of several different kinds of determinism, and the free will/determinism problem we will be discussing is one of a family of related problems. For instance, *logical* determinism is the thesis that the principle of bivalence holds for all propositions, including propositions about our future actions, and the problem of free will and logical determinism is the problem of understanding how, if at all, we can have free will if there are truths about what we will do tomorrow. *Theological* determinism is the thesis that God exists and has infallible knowledge of all true propositions including propositions about our future actions; the problem of free will and theological determinism is the problem of understanding how, if at all, we can have free will if God (who *cannot* be mistaken) knows what we are going to do. (For more on logical determinism, see the entry on [fatalism](#), Taylor 1967, Lewis 1976, Sobel 1998, and Westphal 2006. For more on theological determinism, see the entries on [fatalism](#) and [divine foreknowledge](#).) Although there are instructive comparisons to be made concerning logical, theological, and nomological determinism, these are outside the scope of this article. (For some comparisons between arguments for incompatibilism and arguments for logical fatalism, see van Inwagen 1983, Mackie 2003, Perry 2004, and Vihvelin 2008.) We will be restricting ourselves to nomological determinism, and to arguments for the claim that free will is incompatible with nomological determinism.

At a first approximation, nomological determinism (henceforth “determinism”), is a contingent and empirical claim about the laws of nature: that they are deterministic rather than probabilistic, and that they are all-encompassing rather than limited in scope. At a second approximation, laws are deterministic if they entail exceptionless regularities; e.g., that all *F*'s are *G*'s, that all *ABCD*'s are *E*'s, and so on. At a third approximation, the fundamental laws of nature are probabilistic if they say that *F*'s have an objective chance *N* (less than 1) of being *G*'s. (Note that so-called “statistical laws” need not be probabilistic laws; see Armstrong 1983, Loewer 1996a.) The laws of nature are all-encompassing if deterministic or probabilistic laws apply to everything in the universe, without any exceptions. If, on the other hand, some individuals or some parts of some individuals (e.g., the nonphysical minds of human beings) or some of the behaviors of some of the individuals (e.g., the free actions of human beings) do not fall under *either* deterministic or probabilistic laws, then the laws are not all-encompassing.

Given these rough definitions of the difference between deterministic laws, probabilistic laws, and limited laws, we can understand determinism as the thesis that a complete description of the state of the world at any time *t* and a complete statement of the laws of nature together entail every truth about the world at every time later than *t*. Alternatively, and using the language of possible worlds: Determinism is true at a possible world *w* iff the following is true at that world: Any world which has the same laws of nature as *w* and which is exactly like *w* at any time *t* is

exactly like  $w$  at all times which are future relative to  $t$ . (See van Inwagen 1983, Earman 1986, Ginet 1990, and the entry on [causal determinism](#).)

Let's call a world "deterministic" iff the thesis of determinism is true at that world; non-deterministic iff the thesis of determinism is false at that world. There are two very different ways in which a world might be non-deterministic. A world might be non-deterministic because at least some of its fundamental laws are probabilistic, or a world might be non-deterministic because it has no laws or because its laws are not all-encompassing. Let's call worlds which are non-deterministic in *only* the first way "probabilistic worlds" and let's call worlds which are non-deterministic in the second way "lawless" or "partly lawless" worlds.

Determinism is a thesis about the laws, but it should not be confused with a philosophical analysis or account of lawhood. In particular, determinism should not be confused with the view of laws that has been called "the governing conception of laws" (Beebe 2000), "the pushy explainer view" (entry on [causal determinism](#)), and, most commonly, "the necessitarian view". It's easy to get confused because determinism is often formulated in a loose and misleading way, e.g. as the thesis that facts about the past metaphysically determine or necessitate or "fix" all future facts. Determinism is a claim about the relation of entailment that holds between, on the one hand, statements of law and statements of particular fact at a time, and, on the other hand, statements of particular fact at any later time. This claim about entailment relations is neutral between different accounts of lawhood, ranging from the so-called "naïve regularity" account (Swartz 1986) to broadly Humean or "best system" accounts (Lewis 1973, Earman 1986, Loewer 1996a, Beebe 2000, Schaffer 2008) to various kinds of necessitarian accounts (Shoemaker 1980, Armstrong 1983, Carroll 1994 and 2008).

Note that determinism is not a thesis about predictability. Determinism is a thesis about the statements of law that correctly describe our world; it says nothing about whether these statements are knowable by finite beings, let alone whether they could, even in principle, be used to predict all future events. Chaos theory tells us that some deterministic systems are very difficult to predict. Quantum mechanics tells us, at least according to some interpretations, that the behavior of probabilistic systems is, in some respects, easy to predict. And it is at least arguable that the behavior of a perfectly rational but lawless or partly lawless creature is highly predictable.

Note that determinism is not a thesis about causation; in particular, it is not the thesis that every event has a cause. If, as many people now believe, the fundamental laws are probabilistic rather than deterministic, this doesn't mean that there is no causation; it just means that we have to revise our theories of causation to fit the facts. And this is what philosophers of causation have done; there are probabilistic versions of "covering law" theories of causation, of counterfactual theories of causation, and so on, for all major theories of causation. (See entries on the [metaphysics of causation](#) and [counterfactual theories of causation](#).) It is now generally accepted that it might be true that every event has a cause even though determinism is false and thus some events lack sufficient or deterministic causes.

More controversially, it might be true that every event has a cause even if our world is neither deterministic nor probabilistic. If there can be causes without laws (if a particular event, object,

or person can be a cause, for instance, without instantiating a law), then it might be true, even at a lawless or partly lawless world, that every event has a cause (Anscombe 1981, van Inwagen 1983).

It's less clear whether determinism entails the thesis that every event has a cause. Whether it does so or not depends on what the correct theory of causation is; in particular, it depends on what the correct theory says about the relation between causation and law.

What is clear, however, is that we should *not* make the assumption, almost universally made in the older literature, that the thesis, that every event has a cause, is equivalent to the thesis of determinism. This is an important point, because some of the older arguments in the literature *against* incompatibilism assume that the two claims are equivalent. (For instance, the argument that an event that is not causally determined is random, and for that reason not under anyone's control. Hobart 1934, Ayer 1954.)

In the older literature, it was assumed that determinism is the working hypothesis of science, and that to reject determinism is to be against science. This no longer seems plausible. Some people think that quantum mechanics has shown determinism to be false. This remains controversial (Loewer 1996b), but it is now generally agreed that we can reject determinism without accepting the view that the behavior of human beings falls outside the scope of natural laws. If mechanism is the thesis that human behavior can be explained in the same kind of way — in terms of events, natural processes, and laws of nature — as everything else in the universe, then we can reject determinism without rejecting mechanism.

Note, finally, that determinism neither entails physicalism nor is entailed by it. There are possible worlds where determinism is true and physicalism false; e.g., worlds where minds are nonphysical things which nevertheless obey deterministic laws (van Inwagen 1998). And there are possible worlds (perhaps our own) where physicalism is true and determinism is false.

So much for determinism. What about free will? Here it is very difficult to say anything without saying something that will be contested by some philosopher. It is generally agreed that problems of free will are problems about our capacity or ability or power to perform certain kinds of actions — actions with the property of being done freely (with free will, of our own free will). It is also generally agreed that we have this capacity only if we at least sometimes exercise it, and so challenges to the existence of free will or to its compatibility with determinism are usually challenges to the claim that anyone (or anyone at a deterministic world) ever acts with free will or acts freely. (I will be using these two expressions interchangeably.) It is also widely agreed that the existence of free will is a necessary condition of the existence of moral responsibility; if determinism has the consequence that we don't have free will, it also has the consequence that we are never morally responsible. But beyond this there is very little agreement about anything. So we will not attempt to provide even a rudimentary definition of free will.

It will be useful, however, to have a precise statement of the thesis at issue between compatibilists and incompatibilists.

Let's define the *free will thesis* as the thesis that at least one non-godlike creature has free will. The free will thesis is a minimal claim about free will; it would be true if one person in the universe acted with free will on one occasion. We won't assume that the free will thesis is true or even possibly true, but let a *free will world* be any possible world where the free will thesis is true. Since non-determinism is the negation of determinism, and since determinism is a contingent thesis, we can divide the set of possible worlds into two non-overlapping subsets: deterministic worlds and non-deterministic worlds.

Given this apparatus, we could define incompatibilism and compatibilism in the following way: incompatibilism is the thesis that no deterministic world is a free will world. (Equivalently, incompatibilism is the claim that necessarily, if determinism is true, then the free will thesis is false.) And we could define compatibilism as the denial of incompatibilism; that is, as the claim that some deterministic worlds are free will worlds. (Equivalently, compatibilism is the claim that possibly, determinism and the free will thesis are both true.)

This way of defining compatibilism is unproblematic. There are compatibilists who are agnostic about the truth or falsity of determinism, so a compatibilist need not be a soft determinist (someone who believes that it is in fact the case that determinism is true and we have free will). But all compatibilists believe that it is at least possible that determinism is true and we have free will. So all compatibilists are committed to the claim that there are deterministic worlds that are free will worlds.

But this definition of incompatibilism has a surprising consequence. Suppose, as some philosophers have argued, that we lack free will because free will is conceptually or metaphysically impossible, at least for nongodlike creatures like us (C.D. Broad 1967, G. Strawson 1986, 1994, 2002). If these philosophers are right, there are no free will worlds. And if there are no free will worlds, it follows that there are no *deterministic* free will worlds. So if free will is conceptually or metaphysically impossible, at least for creatures like us, it follows that incompatibilism (as we have just defined it) is true. But this doesn't seem right. If it is conceptually or metaphysically impossible for us to have free will, then we lack free will *regardless* of whether determinism is true or false. And if that is so, then the incompatibilist cannot say the kind of things she has traditionally wanted to say: that the truth or falsity of determinism is *relevant* to the question of whether or not we have free will, that if determinism were true, then we would lack free will *because* determinism is true, and so on.

If we want to avoid this counter-intuitive result, there is a remedy. Instead of understanding compatibilism and incompatibilism as propositions that are contradictories, we can understand them as propositions that are contraries. That is, we can understand compatibilism and incompatibilism as claims that can't both be true, but that can both be false. Compatibilism and incompatibilism are both false if a third claim, impossibilism, is true. Impossibilism is the thesis that free will is conceptually or metaphysically impossible for non-godlike creatures like us.

If we accept this three-fold classification, we can define our terms as follows: Impossibilism is the thesis that there are no free will worlds. Incompatibilism is the thesis that there are free will worlds but no deterministic world is a free will world. Compatibilism is the thesis that there are free will worlds and free will worlds include deterministic worlds.

The term ‘impossibilism’ is being coined here; however, the position it describes is recognized in the literature under a variety of names: the “no free will either way” view, “non-realism”, “illusionism”. Theorists who defend impossibilism include Double 1991, 1992; Strawson 1986, 1994, and 2002; and Smilansky 1993, 2000. Another kind of impossibilist is the logical fatalist (Taylor 1992).

In the older literature, there were just two kinds of incompatibilists — hard determinists and libertarians. A hard determinist is an incompatibilist who believes that determinism is in fact true (or, perhaps, that it is close enough to being true so far as we are concerned, in the ways relevant to free will) and because of this we lack free will (Holbach 1770, Wegner 2003). A libertarian is an incompatibilist who believes that we in fact have free will and this entails that determinism is false, in the right kind of way (van Inwagen 1983). Traditionally, libertarians have believed that “the right kind of way” requires that agents have a special and mysterious causal power not had by anything else in nature: a godlike power to be an uncaused cause of changes in the world (Chisholm 1964). Libertarians who hold this view are committed, it seems, to the claim that free will is possible only at worlds that are at least partly lawless, and that our world is such a world. (But see Clarke 2003 and O'Connor 2000.) But in the contemporary literature there are incompatibilists who avoid such risky metaphysical claims by arguing that free will is possible at worlds where some of our actions have indeterministic event causes (Kane 1996, Balaguer 2004) or that free will is possible at worlds where some of our actions are uncaused (Ginet 1990). Note that none of these three kinds of incompatibilists (agent-causation theorists, indeterministic event-causation theorists, non-causal theorists) need be libertarians. They may reserve judgment about the truth or falsity of determinism and therefore reserve judgment about whether or not we in fact have free will. They might also be hard determinists because they believe that determinism is in fact true. But what they do believe — what makes them incompatibilists — is that it is *possible* for us to have free will and that our having free will depends on a *contingent fact* about the *laws* that govern the universe: that they are indeterministic in the right kind of way. (See the entry on [incompatibilist theories of free will](#).)

Given these definitions and distinctions, we can now take the first step towards clarifying the disagreement between compatibilists and incompatibilists. Both sides agree that it is conceptually and metaphysically possible for us to have free will; their disagreement is about whether any of the possible worlds where we have free will are deterministic worlds. The compatibilist says ‘yes’; the incompatibilist says ‘no’. Arguments for incompatibilism must, then, be arguments for the claim that necessarily, if determinism is true, we lack the free will we might otherwise have.

## 2. Classification of Arguments for Incompatibilism

A common first response to determinism is to think that it means that our choices make no difference to anything that happens because earlier causes have *predetermined* or “fixed” our entire future. On this view, determinism implies that we have a destiny or fate that we cannot avoid, no matter what we choose or decide and no matter how hard we try.

Man, when running over, frequently without his own knowledge, frequently in spite of himself, the route which nature has marked out for him, resembles a swimmer who is obliged to follow

the current that carries him along; he believes himself a free agent because he sometimes consents, sometimes does not consent, to glide with the stream, which, notwithstanding, always hurries him forward. (Holbach 1770, p. 197. See also Wegner 2003.)

It is widely agreed, by incompatibilists as well as compatibilists, that this is a mistake. Determinism might imply that our choices and efforts have earlier sufficient causes; it does not imply that we don't make choices or that our choices and efforts are causally impotent. Determinism is consistent with the fact that our deliberation, choices and efforts are part of the causal process whereby our bodies move and cause further effects in the world. And a cause is the kind of thing which “make a difference”. If I raise my hand because I chose to do so, then it's true, *ceteris paribus*, that if my choice had not occurred, my hand-raising would not have occurred.

Putting aside this worry, we may classify arguments for incompatibilism as falling into one of two main categories:

1. Arguments for the claim that determinism makes it impossible for us to *cause and control our actions in the right kind of way*.
2. Arguments for the claim that determinism deprives us of the *power or ability to do or choose otherwise*.

Arguments of the first kind focus on the notions of self, causation, and responsibility; the worry is that determinism rules out the kind of causation that we invoke when we attribute actions to persons (“It was Suzy who broke the vase”) and make judgments of moral responsibility. (“It wasn't her fault; Billy pushed her.”) Someone who argues for incompatibilism in this way may concede that the truth of determinism is consistent with the causal efficacy of our deliberation, choices, and attempts to act. But, she insists, determinism implies that the *only* sense in which we are responsible for our actions is the sense in which a chess-playing computer is responsible for its moves. Moral responsibility requires something more than this, she believes. Moral responsibility requires autonomy or self-determination: that our actions are caused and controlled by, and only by, our selves. To use a slogan popular in the literature: We act freely and are morally responsible only if we are the ultimate source of our actions.

Each of us, when we act, is a prime mover unmoved. In doing what we do, we cause certain events to happen, and nothing — or no one — causes us to cause these events to happen. (Chisholm 1964, p. 32)

Free will...is the power of agents to be the ultimate creators or originators and sustainers of their own ends or purposes...when we trace the causal or explanatory chains of action back to their sources in the purposes of free agents, these causal chains must come to an end or terminate in the willings (choices, decisions, or efforts) of the agents, which cause or bring about their purposes. (Kane 1996, p. 4)

Arguments of the second kind focus on the notion of choice. To have a choice, it seems, is to have genuine options or alternatives — different ways in which we can act. The worry is that determinism entails that what we do is the *only* thing we can do, and that because of this we

never *really* have a choice *about anything*, as opposed to being under the (perhaps inescapable) illusion that we have a choice. Someone who argues for incompatibilism in this way may concede that the truth of determinism is consistent with our making choices, at least in the sense in which a chess-playing computer make choices, and consistent also with our choices being causally effective. But, she insists, this is not enough for free will; we have free will only if we have a genuine choice about what actions we perform, and we have a genuine choice only if there is more than one action we *are able* to perform.

A person has free will if he is often in positions like these: he must now speak or be silent, and he *can* now speak and *can* now remain silent; he must attempt to rescue a drowning child or else go for help, and he is *able* to attempt to rescue the child and *able* to go for help; he must now resign his chairmanship or else lie to the members; and he has it within his power to resign and he has it within his power to lie. (van Inwagen 1983, p. 8)

By freedom of the will is meant freedom of action. I have *freedom of action* at a given moment if more than one alternative course of action is then *open to me*. Two or more actions are alternatives if it is logically impossible for me to do more than one of them at the same time. Two or more alternatives are open to me at a given moment if which of them I do is entirely up to my choice at that moment. Nothing that exists up to that moment in time stands in the way of my doing next any one of the alternatives (Ginet 1990, p. 90)

One might question whether these are really independent ways of arguing for incompatibilism, for the following reason: We cause and control our actions in the “right kind of way” (let's call it the “self-determining way”) only if we have the power to do or choose other than what we actually do or choose. And if we have the power to do or choose otherwise, then we also have the power (even if we don't always exercise it) to cause and control our actions in the self-determining way. That is, there is just one power at issue, not two. The power to cause and control one's actions in the self-determining way is the *same power* as the power to do or choose otherwise. And if there is just one power, then either determinism entails that we lack this power or it does not.

There is something to this suggestion. It seems that what distinguishes us from a chess playing computer (or an even more sophisticated machine) is something that makes it true that our actions are *up to us* in some way that the actions of even the most intelligent machines are not. And if our actions are up to us in some special way, then it seems reasonable to suppose that it's true *both* that our power to cause and control our actions differs from the causal powers of mere machines *and* that this difference consists in the fact that we have the power to do or at least choose to do more than one action.

However, the claim that the power to do otherwise and the power of self-determination are one and the same is controversial. Harry Frankfurt has argued that a person who is unable to do or even choose otherwise may nevertheless be morally responsible because she causes and control her actions in the right, self-determining kind of way. (Frankfurt 1969). If he is right, then the power to cause one's actions in the self-determining way does not entail the power to do or choose otherwise

Frankfurt's argument has persuaded many, but remains highly controversial. (For a good collection of critical essays, see Widerker and McKenna 2003. See also Lamb 1993 and Vihvelin 2000b.) Since there is nothing like a consensus concerning Frankfurt's claim, we won't assume that the power of self-determination and the power to do or choose otherwise are one and the same (nor will we assume that they are different). Nor will we assume that either of these powers *must* be understood in a way that entails incompatibilism. That is, we won't assume that the power of self-determination entails the falsity of determinism. An influential compatibilist program identifies a person's self with some subset of her motives (e.g., her values or the desires with which she identifies) and then argues that a person's free actions are those actions which have their source in the agent's self thus understood (Frankfurt 1971, 1988; Watson 1975, 2004; Fischer and Ravizza 1998). Nor will we assume that the power to do otherwise entails the falsity of determinism. This might be true, but we need an argument.

Before we turn to a discussion of either of these two main ways of arguing for incompatibilism, let's pause to discuss a third argumentative strategy.

### **3. Arguments based on Intuition**

Perhaps the most common kind of argument for incompatibilism is an argument that appeals primarily to our intuitions. There are many variations on this way of arguing for incompatibilism, but the basic structure of the argument is usually something along these lines:

If determinism is true, then we are like: billiard balls, windup toys, playthings of external forces, puppets, robots, victims of a nefarious neurosurgeon who controls us by directly manipulating the brain states that are the immediate causes of our actions. Billiard balls, ... have no free will. Therefore if determinism is true, we don't have free will.

Intuition-based arguments are inconclusive. Even if determinism entails that there is *something* we have in common with things which lack free will, it doesn't follow that there are no relevant differences. Billiard balls, toys, puppets, and simple robots lack minds, and having a mind is a necessary condition of having free will. And determinism doesn't have the consequence that all our actions are caused by irresistible desires that are, like the neurosurgeon's direct manipulations, imposed on us by external forces outside our control. (For criticism of "intuition pumps", see Dennett 1984. For discussion of more complicated cases involving indirect manipulation, see Section 3.2 below.)

With this caveat in mind, let's take a closer look at the two most influential intuition-based arguments.

#### **3.1 Garden of Forking Paths Argument**

The Garden of Forking Paths argument (van Inwagen 1993, Fischer 1994, Ekstrom 2000) begins by appealing to the idea that whenever we make a choice we are doing (or think we are doing) something like what a traveler does when faced with a choice between different roads. The only roads the traveler is able to choose are roads which are a continuation of the road he is already on. By analogy, the only choices we are able to make are choices which are a continuation of the

actual past and consistent with the laws of nature. If determinism is false, then making choices really is like this: one “road” (the past) behind us, two or more different “roads” (future actions consistent with the laws) in front of us. But if determinism is true, then our journey through life is like traveling (in one direction only) on a road which has no branches. There are other roads, leading to other destinations; if we could get to one of these other roads, we could reach a different destination. But we can't get to any of these other roads from the road we are *actually* on. So if determinism is true, our actual future is our only possible future; we can never choose or do anything other than what we actually do. (See also Flint 1987 and Warfield 2003 for discussion of a related argument that appeals to the metaphor of our freedom to “add” to the list of truths about the world.)

This is a powerful intuition pump, since it's natural to think of our future as being “open” in the branching way suggested by the road analogy and to associate this kind of branching structure with freedom of choice. But several crucial assumptions have been smuggled into this picture: assumptions about time and causation and assumptions about possibility. The assumptions about time and causation needed to make the analogy work include the following: that we “move” through time in something like the way that we move down a road, that our “movement” is necessarily in one direction only, from past to future, that the past is *necessarily* “fixed” or beyond our control in some way that the future is not. These assumptions are all controversial; on some theories of time and causation (the four-dimensionalist or eternalist theory of time, a theory of causation that permits time travel and backwards causation), they are all false (Lewis 1976, Horwich 1987, Sider 2001, Hoefer 2002).

The assumption about possibility is that possible worlds are concrete spatio-temporal things (in the way that roads are) and that worlds can overlap (literally share a common part) in the way that roads can overlap. But most possible worlds theorists reject the first assumption and nearly everyone rejects the second assumption (Adams 1974, Lewis 1986).

Determinism (without these additional and controversial assumptions) does *not* have the consequence that our “journey” through life is like moving down a road; the contrast between determinism and non-determinism is *not* the contrast between traveling on a branching road and traveling on a road with no branches.

As an argument for incompatibilism, the appeal to the metaphor of the Garden of Forking Paths fails. If the intuitions to which it appeals nevertheless continue to engage us, it is because we think that our range of possible choices is constrained by two factors: the laws and the past. We can't change or break the laws; we cannot causally affect the past. (Even if backwards causation is logically possible, it is not within *our* power.) These beliefs — about the laws and the past — are the basis of the most influential contemporary argument for incompatibilism: the Consequence argument. More of this later.

### **3.2 Manipulation and Design Arguments**

We turn now to a family of arguments that work by appealing to our intuitive response to cases involving two persons, whom we may call “Designer” and “Tool”. Designer designs Tool (in some of the stories, in the way the maker of a robot designs his robot or a god creates a human

being; in other stories, by employing techniques of behavioral engineering). Designer's purpose is to ensure either that Tool performs a specific action (Mele 1995 and 2006, Pereboom 2001, Rosen 2002) or that he will have the kind of psychology and motivational structure which will ensure or make probable that he performs certain kinds of actions and leads a certain kind of life. (See Kane 1996 for discussion of Huxley's *Brave New World* and Skinner's *Walden Two*. See also Pereboom 2001.)

We are supposed to have the intuition that Tool is not responsible because he acts unfreely and that he acts unfreely because of Designer's role in the causation of his actions. Tool performs the actions he performs because that's what he was made (meant, built, designed) to do, and it was Designer who made him that way.

The argument then goes as follows:

1. Tool doesn't act freely and, for that reason, is not morally responsible for what he does.
2. If determinism is true, there is no relevant difference between Tool and *any* normal case of apparently free and morally responsible action.
3. Therefore, if determinism is true no one ever acts freely or is morally responsible for what he does.

We are supposed to accept premise 1 on the grounds of our intuitive response to the story about Tool. The argument for premise 2 is that if determinism is true, then we are like Tool with respect to the fact that we are merely the *proximate* causes of our actions. We do what we do because of our psychological characteristics or 'design', and the causes of our psychological characteristics *ultimately come from outside us, from forces and factors beyond our control*. The only difference between us (in this imagined scenario in which determinism is true) and Tool is that our psychological features are not the causal upshot of the work of a designer (or design team) who had a specific plan for us. But this fact about the remote causes of our actions — that they are caused by a variety of natural causes rather than the intentional acts of a single designer or design team — is not relevant to questions about our freedom and responsibility. Or so it is argued, by the advocates of Manipulation and Design arguments.

The success of an instance of this argument depends on the case used to motivate and justify premise 1. What's required is a case (or set of cases) which both ground our belief that Tool is unfree and which supports premise 2. If there is no such case (or set of cases), the argument fails.

Let's begin with cases of the *Brave New World* variety — cases where children are subjected to intensive behavioral engineering from birth, in a way intended to make them accept their assigned roles in a rigidly hierarchical society. Everything depends on the details, but it is surely not implausible to think that the subjects of *Brave New World* cases lack freedom because their cognitive, evaluational, and volitional capacities have been stunted or impaired in certain ways: “they are incapable of effectively envisaging or seeing the significance of certain alternatives, of reflecting on themselves and on the origins of their motivations”. (Watson 1987.) (See also Wolf 1990, who argues, on the basis of similar cases, that the ability that grounds our freedom and responsibility is the unimpaired capacity to choose and act in accordance with “the True and the Good”.) Determinism does not have the consequence that everyone's cognitive, evaluational and

volitional capacities are impaired in these sorts of ways. So if the Watson/Wolf diagnosis is right, cases of the *Brave New World* variety cannot be used to support premise 2. (For other criticisms of some of the cases, see Haji 1998, Kapitan 2000, and Fischer 2004.)

Defenders of Manipulation and Design arguments claim, however, that the argument works even if these cases are set aside. They say that the intuitive force of the argument depends only on the historical fact that deterministically caused actions are ultimately caused, as are Tool's, by factors and forces outside the agent's control. They say that the argument succeeds even in cases where Designer designs a creature with unimpaired capacities and a 'normal' psychology, perhaps the kind we'd like our children to have, perhaps a rationally egoistic psychology of a kind we would prefer our children not to have (cf. Plum, as described by Pereboom 2001 and Ernie from Mele 2006).

To remind ourselves that we are now considering only a subset of our original cases, let's call the person in these cases 'Normal Tool'. If we met Normal Tool, not knowing the peculiar facts about his history, we would think he was one of us. Pick your favorite example of a person you think is free and responsible; call him "Norm". You may think of Normal Tool as someone who is, for all practical purposes, a psychological and physical duplicate of Norm. (Neither a psychologist nor close friend would be able to tell the difference between them.) If historical facts are part of your account of freedom and responsibility (eg. if you think that a free and responsible agent is someone whose current values and deliberative methods have evolved in certain kinds of ways in response to past experience and critical reflection) you may add these features so that Normal Tool resembles Norm in these ways as well. Despite this, the advocates of the Manipulation and Design argument claim that, when we add to our story about Normal Tool the fact that the causes of his actions trace back, via his 'design' and the relevant facts about the evolution of his 'design', to Designer, our intuitive response is that he is neither free nor responsible.

It is far from obvious that this empirical claim about "our" intuitions is correct. It is a notorious fact that the intuitions that people have in response to philosophical thought experiments can differ a great deal and that the way in which the thought experiment is presented can affect the intuition produced. McKenna (2008) argues that the way in which the stories are told makes all the difference: If the stories are told carefully enough, and in a different order (beginning with the case of a normal deterministic agent who closely resembles Norm) we, or at least many of us, have the intuition that Normal Tool is both free and responsible. If so, the argument fails at its first step. (See also Mele 2006, pp. 139-144, for a different kind of empirical criticism of Pereboom's version of the argument.)

Let's concede, for the sake of argument, that the defenders of the argument are right; we (or most of us, or enough of us) have the intuition that Normal Tool is neither free nor morally responsible, regardless of how the cases are presented. But *why* do we have this intuition? Defenders of the argument say that we have it because we accept a generalization - that deterministic causation by forces outside our control robs us of freedom and responsibility - and Normal Tool's case is an instance of this generalization. But there is a *simpler and narrower explanation* that accounts for our intuition. We might believe that Normal Tool is unfree because he is Designer's creature or "tool"; his entire existence is a means to the achievement of

Designer's ends. Of course, Normal Tool is not *literally* a tool insofar as he has his own purposes, his own values, goals, and plans. But he has these purposes only because Designer gave them to him, and Designer gave him these purposes (and other parts of his nature) so that he would thereby serve Designer's purposes. We believe that it is Designer *rather than Tool* who is the true agent of Tool's actions. In short, we believe that Designer is not just Tool's designer, nor is he just someone who causes Tool, somehow or other, to do things. We believe that Designer is Normal Tool's *controller*; he has, and exercises, the power to cause Normal Tool to do exactly the things he wants him to do. (See also Berofsky 2006.)

If this is the correct explanation of our intuitions concerning Normal Tool's unfreedom, then what explains our intuitions does not also generalize in a way that applies to every case of action at a deterministic world. Causation is not the same as control, and determinism does not have the consequence that every person has a controller, nor does it have the consequence that every action is controlled by earlier events, forces, or factors (Dennett 1984).

It might be objected that the existence of a designer/controller is a *bad reason* for believing that Normal Tool is unfree. What matters, so far as questions about freedom and responsibility are concerned, are *facts about the agent*, not facts about someone else. The existence of a designer/controller is relevant only insofar as it has the consequence that Normal Tool *lacks the kind of control* that is required for freedom and responsibility. ("My freedom to dance is equally impaired whether my legs are paralysed by organic disease or shackled by human hands". Watson 1987, p. 171.)

This may be right. But it is beside the point. A causal explanation is not necessarily a justification, and what causally explains our intuitions about Normal Tool's unfreedom need not also justify those intuitions. It's an empirical question whether this proposed causal explanation of our intuitions is correct. But if it is, then the Manipulation and Design argument is in trouble. For either the explanation is also a justification, in which case there is a relevant difference between Normal Tool and a normal deterministic agent. Or the explanation is not a justification, in which case our intuitions about Normal Tool are suspect because they are based on bad reasons.

A defender of the argument might respond as follows: "Granted, our intuitions may be based on bad reasons. But they might nevertheless direct us towards good reasons. Before we thought about determinism, we thought that Normal Tool is unfree because of the existence of Designer/Controller. This was a bad reason for thinking him unfree. But now that we have reflected about the nature and implications of determinism, we see that Normal Tool is exactly like a normal deterministic agent in a way that is intuitively relevant to freedom and responsibility: the actions of both are ultimately caused by something beyond their control (another agent, in Normal Tool's case; natural forces and factors, in the deterministic case). Therefore, there is no relevant difference between the two and, since Normal Tool acts unfreely, so does the normal deterministic agent."

There are two problems with this reply. First, our intuitions about the case of Normal Tool were supposed to support the verdict of unfreedom about the normal deterministic case. If we can no longer assume that our intuitions are justified, we can no longer use these intuitions to support

our belief about the unfreedom of the normal deterministic agent. We can agree that there are no relevant differences between the two, but this leaves open the possibility that both are free, as well as the possibility that both are unfree. Second, the relevant similarity claimed between Normal Tool and a normal deterministic agent is a *historical fact about the remote causes of their actions*. But historical facts (whether about controllers or other remote causes) are relevant to an agent's freedom only insofar as these facts have the consequence that the agent lacks the kind of control over himself and his actions that is required for freedom and responsibility. The free will/determinism problem is the problem of deciding whether facts about the deterministic causes of actions are the kind of facts that have this consequence. Let us turn, then, to the two main kinds of argument for this claim.

## 4. Self-Determination and the Causal Chain argument

What has this boy to do with it? He was not his own father; he was not his own mother; he was not his own grandparents. All of this was handed to him. He did not surround himself with governesses and wealth. He did not make himself. And yet he is to be compelled to pay. (Darrow 1924, p. 65)

Libertarians and incompatibilists do not want indeterminism for its own sake...indeterminism is something of a nuisance for them. It gets in the way and creates all sorts of trouble. What they want is ultimate responsibility and ultimate responsibility requires indeterminism. (Kane 1989, p. 121)

Let's turn now to arguments for incompatibilism based on the idea that a free and responsible action is an action that is caused and controlled by its agent in a special self-determining way, a way that is incompatible with deterministic event-causation.

The most popular instance of this kind of argument is an argument that I will call "the desperate defense attorney's argument" (Darrow 1924). The defense attorney's argument is simple:

1. My client is responsible for his crime only if he "made himself" — that is, only if *he* caused himself to be the kind of person he is.
2. My client did not make himself.
3. Therefore my client is not responsible for his crime.

The defense attorney is trying to persuade the jurors that his client is not responsible for his action, but not for any of the standard excusing conditions — insanity, accident, mistaken belief, duress, mental handicap, and so on. Nor does he claim that there is anything that distinguishes his client from any of the rest of us. His argument is that his client is not responsible because he did not make himself. But none of us has made ourselves (at least not from scratch) — we are all the products of heredity and environment. So if we accept the defense attorney's argument, it appears that we are committed to the conclusion that *no one is ever responsible for anything*.

It's not clear that this is an argument for incompatibilism. It's an argument for incompatibilism only if it's an argument for hard determinism — that is, if it's an argument for the thesis that

*determinism* is true and *because of this* we are never responsible for anything. Let's take a closer look.

What's the argument for premise (2)? After all, we do make our selves, at least in the garden-variety way in which we make other things: we plant gardens, cook dinners, build boats, write books and, over the course of our lives, re-invent, re-create, and otherwise “make something of ourselves”. Of course we don't do any of these things “from scratch”, without help from anyone or anything else, but it's impossible (or at least impossible for human beings) to make anything from scratch. The truth or falsity of determinism has no bearing on this point.

(See G. Strawson 1986, 1994, and 2002 for an argument for the *impossibility* of “true responsibility” that is based on a more sophisticated version of the defense attorney's argument. See also Smilansky 2000. See Clarke 2005 and Mele 1995 for a critique of Strawson's argument.)

If we pressed our defense attorney (or brought in a philosopher to help him out), we might get the following reply: Making our actions (and thereby our selves) in the way that we make dinners, gardens, boats, and books is not good enough because we are talking about moral responsibility. In order to be truly responsible — in a way that allows us to justify blame and punishment — for our actions *we* must be the *ultimate sources* or *first causes* of the choices we make. Of course we can't be the *sole cause* of what we do because of our choice; we have to work with the raw material of our physical bodies and that part of the way we are that is the product of external causes. And of course the success of our attempts to act on our choices depends on factors outside our control (Nagel 1976). But in order to be responsible for the upshots of our choices, we must be responsible for our choices. And we are responsible for our choices only if *we cause our choices and no one and nothing causes us to cause them*.

The defense attorney (or philosopher) is defending premise (2) by arguing for a certain interpretation of premise (1) — that our responsibility for our actions requires that we have “made ourselves” in the sense that, over the course of our lives, we have frequently been the first cause of the choices that result in actions and thus eventually (albeit often in ways we can neither predict nor control) to changes in our selves. In arguing this way, he has shifted the focus of the argument from the *obviously* impossible demand that moral responsibility requires (entirely) self-made selves to the intuitively appealing and at least *not obviously* impossible demand that moral responsibility requires what Robert Kane has called “ultimate responsibility” (that we are the ultimate sources or first causes of at least some of our choices, decisions, intentions, or acts of will). (See Kane 1996. See also Chisolm 1964, Zimmerman 1988, Clarke 1993, 1996, and 2003, O'Connor 1995 and 2000, and Pereboom 2001.)

Insofar as this is an argument for incompatibilism, it is an argument for the claim that determinism entails that we are *mere links in a causal chain* and therefore *merely* the proximate causes of our choices. Our choices cause our actions, but our choices are caused by our beliefs and desires (or values, reasons, character traits, etc.) and these in turn have external causes. So if determinism is true, then the way in which our actions are caused does not, after all, differ in any relevant way, from the way that the ‘actions’ of the chess-playing computer are caused.

This brings us to the philosopher's version of the defense attorney's argument. Let's call it the 'Causal Chain argument'. (For variations on this kind of argument, see Kane 1996, McKenna 2001, and Pereboom 2001.)

1. We have free will (of the kind required for moral responsibility) only if *we* are the ultimate sources (originators, first causes) of our choices.
2. If determinism is true, then everything we do is ultimately caused by events and circumstances outside our control.
3. If everything we do is ultimately caused by events and circumstances outside our control, then *we* are not the ultimate sources (originators, first causes) of our choices.
4. Therefore, if determinism is true, we are not the ultimate sources of our choices.
5. Therefore, if determinism is true, we don't have free will (of the kind required for moral responsibility).

Premise (2) follows from the definition of determinism (at least given two widely accepted assumptions: that there is causation in a deterministic universe and that causation is a transitive relation). (For some doubts about the latter assumption, see Hall 2000, and Hitchcock 2001). Premise (3) is clearly true. So if we want to reject the conclusion, we must reject Premise (1).

Compatibilists have argued against (1) in two different ways. On the positive side, they have argued that we can give a satisfactory account of the (admittedly elusive) notion of self-determination without insisting that self-determination requires us to be the first causes of our choices (see Bok 1998, Dennett 1984, Fischer 1994, Fischer and Ravizza 1998, Frankfurt 1971 and 1988, Wallace 1994 and 2002, Watson 1975, 1987 and 2004, Wolf 1990). On the negative side, compatibilists have challenged (1) by arguing that it is of no help to the incompatibilist: if we accept (1), we are committed to the conclusion that free will and moral responsibility are *impossible*, regardless of whether determinism is true or false.

The challenge to (1) takes the form of a dilemma: Either determinism is true or it's not. If determinism is true, then my choices are ultimately caused by events and conditions outside my control, so I am not their first cause and therefore, if we accept (1), I am neither free nor responsible. If determinism is false, then something that happens inside me (something that I call "my choice" or "my decision") might be the *first event* in a causal chain leading to a sequence of body movements that I call "my action". But since this event is not causally determined, whether or not it happens is a matter of chance or luck. Whether or not it happens has nothing to do with *me*; it is *not under my control* any more than an involuntary knee jerk is under my control. Therefore, if determinism is false, *I* am not the first cause or ultimate source of my choices and, if we accept (1), I am neither free nor responsible (Ayer 1954, Wolf 1990).

In order to defend (1) against the so-called "determined or random" dilemma, above, the incompatibilist has to offer a positive account of the puzzling claim that *persons* are the first causes of their choices. The traditional incompatibilist answer is that this claim must be taken literally, at face value. We — agents, persons, enduring things — are causes with a very special property: we initiate causal chains, but nothing and no one causes us to do this. Like God, we are uncaused causers, or first movers. For instance, if Joe deliberately throws a rock, which breaks a window, then the window's breaking (an event) was caused by Joe's throwing the rock (another

event), which was caused by Joe's choice (another event). But Joe's choice was not caused by any further event, not even the event of Joe's thinking it might be fun to throw the rock; it was caused by *Joe himself*. And since Joe is not an event, he is not the kind of thing which can be caused. (Or so it is argued, by defenders of the conceptual possibility of agent-causation. See Chisolm 1964 and O'Connor 1995 and 2000, and Pereboom 2001.)

Many philosophers think that agent-causation is either incoherent or impossible, due to considerations about causation. What sense does it make to say that a person or other enduring thing, as opposed to a *change* in a thing, or the *state* of a thing at a time, is a cause? (Broad 1952, Bok 1998. See also Clarke 2003 for a detailed and sympathetic examination of the metaphysics of agent-causation, which ends with the conclusion that there are, on balance, reasons to think that agent-causation is impossible.)

Others (Broad 1952, Taylor 1960, van Inwagen 2000, Mele 2006) have argued that even if agent-causation is possible, it would not solve the problem of transforming an undetermined event into one which is in *our* control in the way that our free choices must be.

Recently some incompatibilists have responded to the “determined or random” dilemma in a different way: by appealing to the idea of probabilistic causation (Kane 1996). If our choices are events which have probabilistic causes (e.g., our beliefs, desires, and other reasons for acting), then it no longer seems plausible to say that we have *no* control over them. We make choices for reasons, and our reasons cause our choices, albeit indeterministically. Kane's reply may go some way towards avoiding the second (no control) horn of the dilemma. But it doesn't avoid the first horn. If our reasons cause our choices, then our *choices* are not the first causes of our actions. And our reasons are presumably caused, either deterministically or probabilistically, so they are not the first causes of our actions either. But then our actions are ultimately caused by earlier events over which we have no control and we are not the ultimate sources or first causes of our actions.

## 5. Choice and the Consequence Argument

... determinism ... professes that those parts of the universe already laid down absolutely appoint and decree what the other parts shall be. The future has no ambiguous possibilities hidden in the womb. The part that we call the present is compatible with only one totality. Any other future complement than the one fixed from eternity is impossible. The whole is in each and every part, and welds it with the rest into an absolute unity, an iron block, in which there can be no equivocation or shadow of turning... necessity on the one hand and impossibility on the other are the sole categories of the real. Possibilities that fail to get realized, are, for determinism, pure illusions; they never were possibilities at all. (James, 1884, p. 150-51)

We think that we make choices, and we think that our choices typically make a difference to our future. We think that there is a point to deliberation: how we deliberate — what reasons we consider — makes a difference to what we choose and thus to what we do. We also think that when we deliberate we are trying to decide which, of many possible futures, is the one we want to make actual. That is, we believe that there *really* is more than one choice we *can* make, more

than one action we *can* perform, and more than one future which is, at least partly, *within our power* to bring about.

Our beliefs about our power with respect to the future contrast sharply with our beliefs about our lack of power with respect to the past. We don't think we have *any* choice about the past. We don't deliberate about the past; we think it irrational to do anything aimed at trying to change or affect the past. ("You had your chance; you blew it. It's too late now to do anything about it.") Our beliefs about our options, opportunities, alternatives, possibilities, abilities, powers, and so on, are all future-directed. We may summarize this contrast by saying that we think that the future is "open" *in some sense* that contrasts with the non-openness or "fixity" of the past.

Although we don't think we (now) have a choice about the past, we have beliefs about what *was* possible for us in the past. When called upon to defend what we did, or when we blame or reproach ourselves, or simply wonder whether we did the right thing (or the sensible thing, the rational thing, and so on), we evaluate our action by comparing it to what we believe were our other possible actions, *at that time*. We blame, criticise, reproach, regret, and so on, only insofar as we believe we had alternatives. And if we later discover that we were mistaken in believing that some action X was among our alternatives, we think it is irrational to criticise or regret our failure to do X.

Is determinism compatible with the truth of these beliefs? In particular, is it compatible with the belief that we are often able to choose and do more than one action?

Incompatibilists have traditionally said "No". And it's not hard to see why. If we think of 'can' in the "open future" way suggested by the commonsense view, then it's tempting to think that the past is *necessary* in some absolute sense. And it's natural to think that we are able to do otherwise only if we *can do otherwise given the past*; that is, only if our doing otherwise is a possible continuation of the past. If we follow this train of thought, we will conclude that we are able to do otherwise only if our doing otherwise is a *possible continuation of the past consistent with the laws*. But if determinism is true, there is only *one* possible continuation of the past consistent with the laws. And thus we get James' conclusion: Determinism says that the actual future is necessary and any other future is impossible. What will be, *must* be. What will not be, *cannot* be.

But this argument is too quick. There is an alternative explanation for our beliefs about the "open" future as opposed to the "fixed" past — the direction of causation. Causal chains run from past to future, and not in the other direction. Our deliberation causes our choices, which cause our actions. But not the other way around. Our choices cause future events; they never cause past events. Why causation works this way is a deep and difficult question, but the leading view, among philosophers of science, is that the temporal asymmetry of causation is a fundamental but contingent fact about our universe. If things were different enough — if we could travel backwards in time — then we would have an ability that we don't actually have — the ability to causally affect past persons and things (Horwich 1987, Lewis 1976). If this is right, then we don't need to suppose that the past is metaphysically or absolutely necessary in order to explain the open future/fixed past contrast. The past *could have been different*. But, given the

way things actually are (given the contingent fact that accounts for the forward direction of causation), there is nothing that *we are able to do that would cause the past to be different*.

This alternative explanation of our commonsense belief about the contrast between open future and fixed past allows the compatibilist to say the kind of things that compatibilists have traditionally wanted to say: The “can” of freedom of choice is the ‘can’ of causal and counterfactual dependence. Our future is open because it depends causally and counterfactually, on our choices, which in turn depend, causally and counterfactually, on our reasons. (At least in the normal case, where there is neither external constraint nor internal compulsion or other pathology.) If our reasons were different (in the appropriate way), we would choose otherwise, and if we choose otherwise, we would do otherwise (Moore 1912, Lehrer 1980). And our reasons *can* be different, at least in the sense that we, unlike simpler creatures and young children, have the ability (skill, capacity) to critically evaluate the reasons (beliefs, desires, values, principles, and so on) that we have and, at least sometimes exercise, the ability to change our reasons (Bok 1998, Dennett 1984, Fischer 1994, Frankfurt 1988, Lehrer 1980, Wallace 1994, Watson 1975, 1987, and 2004, Wolf 1990). All this is compatible with determinism. So the truth of determinism is compatible with the truth of our commonsense belief that we really do have a choice about the future, that we really can choose and do other than what we actually do.

Incompatibilists think that this, and *any* compatibilist account of the ‘can’ of freedom of choice, is, and *must be*, mistaken. But they have traditionally had a hard time explaining why. The Consequence Argument, due chiefly to Ginet and van Inwagen (Ginet 1966, 1980, 1983, 1990, van Inwagen 1974, 1975, 1983; but see also Wiggins 1973 and Lamb 1977) is widely regarded as the best argument for the conclusion that if determinism is true, then no one ever *really* has a choice about anything. In the remainder of this section we will take a closer look at van Inwagen's version of this important and influential argument.

In *An Essay on Free Will* (1983), van Inwagen presents three formal arguments which, he says, are intended as three versions of the same basic argument, which he characterized as follows:

If determinism is true, then our acts are the consequence of laws of nature and events in the remote past. But it's not up to us what went on before we were born, and neither is it up to us what the laws of nature are. Therefore, the consequences of these things (including our present acts) are not up to us. (p. 56)

We will begin by looking at the third version of the argument (the modal argument) and conclude by considering David Lewis's criticism of the first (truth-functional) version of the argument. The modal argument uses a modal sentential operator which van Inwagen defines as follows: ‘**N***p*’ abbreviates ‘*p* and no one has, or ever had, any choice about whether *p*’. Van Inwagen tells us that the logic of ‘**N**’ includes these two inference rules, where  $\Box p$  asserts that it is logically necessity that *p*:

**Alpha:**

From  $\Box p$ , we may infer **N***p*.

**Beta:**

From  $\mathbf{N}p$  and  $\mathbf{N}(p \supset q)$ , we may infer  $\mathbf{N}q$ .

In the argument below, ' $L$ ' is an abbreviation for a sentence expressing a conjunction of all the laws of nature; ' $H$ ' is a sentence expressing a true proposition about the total state of the world at some time in the distant past before any agents existed; ' $\Box$ ' is 'it is logically necessary that'; ' $\supset$ ' is the material conditional, and ' $P$ ' is a dummy for which we may substitute any sentence which expresses a true proposition.

The argument is a conditional proof: Assume determinism and show that it follows that no one has, or ever had, a choice about *any* true proposition, including propositions about the apparently free actions of human beings.

1.  $\Box((H \& L) \supset P)$  definition of determinism
2.  $\Box(H \supset (L \supset P))$  from 1, by modal and sentential logic
3.  $\mathbf{N}(H \supset (L \supset P))$  from 2, by rule Alpha
4.  $\mathbf{N}H$  premise, fixity of past
5.  $\mathbf{N}(L \supset P)$  from 3, 4, by rule Beta
6.  $\mathbf{N}L$  premise, fixity of laws
7.  $\mathbf{N}P$  from 5, 6, by rule Beta

Premises (1) and (2) follow from determinism. (3) follows from (2), by application of rule Alpha. Rule Alpha is uncontroversial. If something is a logically necessary truth, then no one has, or ever had, any choice about it.

Premises 4 and 6 also look uncontroversial.  $\mathbf{N}$  necessity isn't logical or metaphysical necessity. We can insist that the laws and the distant past *could*, in the broadly logical sense, have been different, so neither  $\Box H$  nor  $\Box L$  are true. But it still seems undeniably true that *we have no choice about* whether the laws and the distant past are the way they are; there is nothing that *we* are able to do that would make it the case that either the laws or the distant past are different from the way they actually are.

Rule Beta is the key to the argument. It's what makes the difference between the modal version of the Consequence Argument and an argument widely agreed to be fallacious.

$\Box(P \supset Q)$   
 $P$   
 Therefore,  $\Box Q$

An example of this invalid inference is an argument sometimes called "the fatalist fallacy":

$\Box(\text{it's true that it will rain tomorrow} \supset \text{it will rain tomorrow})$   
It's true that it will rain tomorrow  
Therefore,  $\Box(\text{it will rain tomorrow})$

Another example:

$\Box((H \ \& \ L) \supset P)$   
 $H \ \& \ L$   
Therefore,  $\Box P$

On the other hand, the following is a valid inference:

$\Box P$   
 $\Box(P \supset Q)$   
Therefore,  $\Box Q$

The necessity expressed by the ‘no choice about’ operator is not logical or metaphysical necessity. But it might nevertheless be similar enough for Beta to be a valid rule of inference. Or so argued van Inwagen, who gave examples:

$\mathbf{N}(\text{The sun explodes in the year 2000})$   
 $\mathbf{N}(\text{The sun explodes in the year 2000} \supset \text{All life on earth ends in the year 2000})$   
Therefore,  $\mathbf{N}(\text{All life on earth ends in the year 2000})$

An early response to the Consequence argument was to argue that Beta is invalid *because* a compatibilist account of ability to do otherwise is correct (Gallois 1977, Narveson 1977, Foley 1979, Slote 1982, Flint 1987). For instance, if “*S* is able to do *X*” means “if *S* chose to do *X*, *S* would do *X*”, then the premises of the argument are true (since even if *S* chose to change the laws or the past, she would not succeed), but the conclusion is false (since determinism is consistent with the truth of conditionals like “if *S* chose to raise her hand, she would”).

Incompatibilists were unmoved by this response, saying, in effect, that the validity of Beta is more plausible than the correctness of any compatibilist account of ability to do otherwise. They pointed out that there was no agreement, even among compatibilists, about how such an account should go, and that the simplest accounts (so-called “Conditional Analyses”, originally proposed by G.E. Moore 1912) had been rejected, even by compatibilists.

(For criticism of Conditional Analyses, see Austin 1961, Berofsky 2002, Lehrer 1968, and van Inwagen 1983. For defense of a compatibilist account of ability to do otherwise, see Kapitan 1991, 1996, and 2002, Lehrer 1980, Bok 1998, Smith 1997 and 2004, Campbell 2004, Perry 2004, and Vihvelin 2004. For defense of the claim that the ability to do otherwise is not necessary for moral responsibility and/or any variety of free will worth wanting, see Frankfurt 1969, Dennett 1984, Fischer 1994, and Fischer and Ravizza 1998.)

More recently, van Inwagen has conceded that Beta is invalid (van Inwagen, 2000). McKay and Johnson (1996) showed that Beta entails Agglomeration:

$\mathbf{N}p$   
 $\mathbf{N}q$   
Therefore,  $\mathbf{N}(p \ \& \ q)$

Agglomeration is uncontroversially invalid. To see this, let 'p' abbreviate "The coin does not land heads", let 'q' abbreviate "The coin does not land tails", and suppose that it's a fair coin which isn't tossed but someone could have tossed it (McKay and Johnson 1996).

(For cases that are counterexamples to Beta, see Widerker 1987, Huemer 2000, and Carlson 2000.)

Van Inwagen proposed to repair the Consequence argument by replacing 'N' with 'N\*', where 'N\*p' says "*p* and no one can, or ever could, do anything such that if she did it, *p* might be false". Agglomeration is valid for 'N\*', and thus this particular objection to the validity of Beta does not apply.

It has also been suggested (Finch and Warfield 1998) that the Consequence argument can be repaired by keeping 'N' and replacing Beta with Beta 2:

**Beta 2:**

From  $\mathbf{N}p$  and  $\Box(p \supset q)$ , we may infer  $\mathbf{N}q$

This would yield the following argument:

$\mathbf{N}(L \ \& \ H)$       fixity of laws and past  
 $\Box((L \ \& \ H) \supset P)$       determinism  
 $\mathbf{N}P$                   from 1, 2 by Beta 2

Other ways of repairing the argument have been proposed by O'Connor 1993 and Huemer 2000.

These revised versions of the Consequence argument may not be as plausible as the original version, but it still looks as though the compatibilist is in trouble. For it seems plausible to suppose that there is nothing that we are able to do that *might* make it the case that either *H* or *L* is false. And it seems plausible to suppose that we have no choice about whether (*H* & *L*). We need to dig deeper to criticize the argument.

Lewis (1981) begins with a clear and elegant statement of the first version of the argument. Paraphrasing slightly, it goes like this:

Suppose that determinism is true, and that I just put my hand down on my desk. As a compatibilist, I claim that this is a free but determined act. I was able to act otherwise, for instance to raise my hand. But there is a true historical proposition *H* about the intrinsic state of the world long ago, and a true proposition *L* specifying the laws of nature, such that *H* and *L* jointly determine what I did, and jointly contradict the proposition that I raised my hand. If I had raised my hand, then at least one of three things would have been true: contradictions would

have been true, *H* would not have been true, or *L* would not have been true. So if I claim that I am able to raise my hand, I am committed to the claim that I have one of three incredible abilities: the ability to make contradictions true, the ability to change the past, or the ability to break (or change) the laws. It's absurd to suppose that I have any of these abilities. Therefore, by reductio, I could not have raised my hand.

Lewis replies by agreeing that if he had raised his hand contradictions would still not be true, and agreeing that if he had raised his hand, *H* would still have been true. What about *L*? Lewis draws a distinction between two counterfactuals:

(C1) If I had raised my hand, *L* would not have been true. (“... a law would have been broken”)

(C2) If I had raised my hand, my act would have been or caused an event which entails not-*L* (“... would have been or caused a law-breaking event”)

Lewis accepts (C1) (for the theory of counterfactuals which supports this claim, see Lewis 1973 and 1979), but denies (C2). He then distinguishes between two ability claims:

(A1) I am able to do something such that if I did it, *L* would not have been true (“... a law would have been broken”)

(A2) I am able to do something such that if I did it, my act would have been or caused an event which entails not-*L* (“... would have been or caused a law-breaking event”)

As a compatibilist who accepts (C1), Lewis is committed to the truth of (A1). But since he rejects (C2), he is not committed to (A2). And it is (A2) rather than (A1) which describes an incredible ability. Or so Lewis argues.

If Lewis is right, the Consequence argument fails to show that any compatibilist account of ability to do otherwise *must* be mistaken. Lewis is basically saying that the Consequence argument equivocates between two ways of understanding ability claims (when they are applied to propositions) and two correspondingly different ways of understanding the “N” necessity of a proposition. The compatibilist is committed to the claim that free determined agents have the ability to do something such that if they did it, then either *H* would not have been true or *L* would not have been true. But this is just another way of saying something that compatibilists have always said: that someone may have the ability to do *X* even though she would not exercise her ability unless different circumstances obtained (e.g., circumstances which provide her with reasons for doing *X*). The compatibilist is *not* committed to the claim that free determined agents have the ability to *cause* either the past or the laws to be different. So the compatibilist is not committed to any incredible claims about the abilities of free agents. (See also Fischer 1983, 1988, and 1994, Horgan 1985, Kapitan 1991, 1996, and 2002, Carlson 2000, and Schneider 2004.)

If the aim of the Consequence argument was to show that no compatibilist account of ‘could have done otherwise’ *can* succeed, then Lewis is surely right; the reductio fails. The distinction between (A1) and (A2) (and a similar distinction concerning ability with respect to the past)

permits the compatibilist to avoid making incredible claims about the powers of free determined agents. On the other hand, the incompatibilist surely has a point when she complains that it is difficult to believe that anyone has the ability described by (A1). We believe that our powers as agents are constrained by the past *and* by the laws. Granted, one way to understand this belief is the way the compatibilist suggests: *we lack causal power* over the *past* and the *laws*. But it's natural to understand the constraint in a different, simpler way: we have the power to do only those things which we can do, *given the past and the laws*. And this leads more or less directly to the incompatibilist conclusion that if determinism is true, then we are never able to do otherwise.

This brings us back to our starting point. Our common sense web of beliefs about ourselves as deliberators, choosers, and agents includes the belief that the future is open in some sense that the past is not. It also includes the belief that our abilities and powers are constrained by the laws. One way of understanding these beliefs leads to incompatibilism; another way does not. Which one is right?

The Consequence argument is an attempt to provide an *argument* in defense of the incompatibilist's way of understanding these common sense beliefs. Even if it fails as a reductio, it has been successful in other ways. It has made it clear that the free will/determinism problem is a metaphysical problem and that the underlying issues concern questions about our abilities and powers, as well as more general questions about the nature of causation, counterfactuals, and laws of nature. Can the abilities or powers of choosers and agents be understood as a kind of natural capacity or disposition? Is there a viable incompatibilist alternative? How should we understand counterfactuals about the alternative actions and choices of agents at deterministic worlds? Is the compatibilist proposal about the way in which the laws and past constrain us defensible? Are incompatibilists committed to the defense of a particular view about the nature of laws of nature? Are they committed to the rejection of a Humean view, for instance?

Insofar as the Consequence argument has pointed us in the direction of these deep and difficult underlying metaphysical questions, it represents a significant step forward in the discussion of one of the most intractable problems of philosophy. (For discussion of some of these modal and metaphysical issues, see Lewis 1973 and 1979, Berofsky 1987, Smith 1997 and 2004, Vihvelin 1990, 2000a and 2004, Stone 1998, Markosian 1999, O'Connor 2000, Beebe and Mele 2002, Beebe 2003, Bennett 2003, Clarke 2003, Perry 2004, and van Inwagen 2004a.)

## Bibliography

- Adams, Robert, 1974, "Theories of Actuality", *Noûs*, 8: 211-31.
- Anscombe, Elizabeth, 1981, "Causality and Determination", in *Metaphysics and the Philosophy of Mind*, Minneapolis: University of Minnesota Press: 133-147.
- Armstrong, David, 1983, *What is a Law of Nature?*, Cambridge: Cambridge University Press.
- Austin, J.L., 1961, "Ifs and Cans", *Philosophical Papers*, London: Oxford University Press: 205-232.
- Ayer, A. J., 1954, "Freedom and Necessity", in his *Philosophical Essays*, New York: St. Martin's Press: 3-20. Reprinted in Pereboom 1997.

- Balaguer, Mark, 2004, “A Coherent, Naturalistic, and Plausible Formulation of Libertarian Free Will”, *Noûs*, 38: 379-406.
- Beebe, Helen, 2000, “The Non-Governing Conception of Laws”, *Philosophy and Phenomenological Research*, 61: 571-594.
- —, 2002, “Reply to Huemer on the Consequence Argument”, *The Philosophical Review*, 111, 235-241.
- —, 2003, “Local Miracle Compatibilism”, *Noûs*, 37: 258-277.
- Beebe, Helen and Mele, Alfred, 2002, “Humean Compatibilism”, *Mind*, 111(442): 201-23.
- Bennett, Jonathan, 2003, *A Philosophical Guide to Conditionals*, Oxford: Clarendon Press.
- Berofsky, Bernard, 1987, *Freedom from Necessity*, London: Routledge & Kegan Paul.
- —, 2002, “Ifs, Cans, and Free Will: The Issues”, in Kane 2002B, 181-201.
- —, 2006, “Global Control and Freedom”, *Philosophical Studies*, 131: 419-445.
- Bok, Hilary, 1998, *Freedom and Responsibility*, Princeton: Princeton University Press.
- Broad, C.D., 1934, *Determinism, Indeterminism, and Libertarianism*, London: Routledge & Kegan Paul.
- Campbell, Joseph, 2005, “Compatibilist Alternatives”, *Canadian Journal of Philosophy*, 35: 387–406.
- Carlson, Erik, 2000, “Incompatibilism and the Transfer of Power Necessity”, *Noûs*, 34: 277-290.
- Carroll, John, 1994, *Laws of Nature*, Cambridge: Cambridge University Press.
- —, 2008, “Nailed to Hume's Cross?”, *Contemporary Debates in Metaphysics*, J. Hawthorne, T. Sider, D. Zimmerman (eds.), Malden, MA: Blackwell Publishing, 67-81.
- Chisolm, Roderick, 1964, “Human Freedom and the Self”, *The Lindley Lectures*, Copyright by the Dept. of Philosophy, University of Kansas. Reprinted in Watson 2003.
- Clarke, Randolph, 1993, “Towards a Credible Agent-Causal Account of Free Will”, *Noûs*, 27: 191-203. Reprinted in Watson 2003.
- —, 1996, “Agent Causation and Event Causation in the Production of Free Action”, *Philosophical Topics*, 24(2): 19-48.
- —, 2003, *Libertarian Accounts of Free Will*, New York: Oxford University Press.
- —, 2005, “On an Argument for the Impossibility of Moral Responsibility”, *Midwest Studies in Philosophy*, 29: 13-24.
- Darrow, Clarence, 1924, *The trial of Leopold and Loeb*, page reference is to “The Trial of Compulsion” in *Attorney for the Damned: Clarence Darrow in the Courtroom*, Arthur Weinberg (ed.), Chicago: The University of Chicago Press, 16-86.
- Dennett, Daniel, 1984, *Elbow Room: The Varieties of Free Will Worth Wanting*, Cambridge, MA: MIT Press.
- Dennett, Daniel & Taylor, Christopher, 2002, “Who's Afraid of Determinism? Rethinking Causes and Possibilities”, in Kane 2002B, 257-277.
- Double, Richard, 1991, *The Non-Reality of Free Will*, New York: Oxford University Press.
- —, 1992, “How Rational Must Free Will Be?”, *Metaphilosophy*, 23: 268-278.
- Earman, John, 1986, *A Primer on Determinism*, Dordrecht: D. Reidel Publishing Company.
- Ekstrom, Laura, 2000, *Free Will: A Philosophical Study*, Boulder, CO: Westview Press.

- Finch, Alicia and Warfield, Ted. A., 1998, "The Mind Argument and Libertarianism", *Mind*, 107(427): 515-28.
- Fischer, John Martin, 1983, "Incompatibilism", *Philosophical Studies*, 43: 121-37.
- —, 1988, "Freedom and Miracles", *Noûs*, 22: 235-252.
- —, 1994, *The Metaphysics of Free Will*, Oxford: Blackwell.
- —, 2004, "Responsibility and Manipulation", *Journal of Ethics*, 8: 145-77.
- Fischer, John Martin and Ravizza, Mark, 1998, *Responsibility and Control: A Theory of Moral Responsibility*, Cambridge: Cambridge University Press.
- Flint, Thomas, 1987, "Compatibilism and the Argument from Unavoidability", *Journal of Philosophy*, 74: 423-40.
- Foley, Richard, 1979, "Compatibilism and Control over the Past", *Analysis*, 39: 70-74.
- Frankfurt, Harry, 1969, "Alternate Possibilities and Moral Responsibility", *Journal of Philosophy*, 66: 820-39. Reprinted in Frankfurt 1988 and Watson 2003.
- —, 1971, "Freedom of the Will and the Concept of a Person", *Journal of Philosophy*, 68: 5-20. Reprinted in Frankfurt 1988 and Watson 2003.
- —, 1988, *The Importance of What we Care About*, Cambridge: Cambridge University Press.
- Gallois, Andre, 1977, "Van Inwagen on Free Will and Determinism", *Philosophical Studies*, 32: 99-105.
- Ginet, Carl, 1966, "Might We Have No Choice?" in *Freedom and Determinism*, K. Lehrer (ed.), New York: Random House, 87-104.
- —, 1980, "The Conditional Analysis of Freedom", in van Inwagen 1980, 171-186.
- —, 1983, "A Defense of Incompatibilism", *Philosophical Studies*, 44: 391-400.
- —, 1990, *On Action*, Cambridge: Cambridge University Press.
- Haji, Ishtiyaque, 1998, *Moral Appraisability*, New York: Oxford University Press.
- Hall, Ned, 2000, "Causation and the Price of Transitivity", *Journal of Philosophy*, 97: 198-222.
- Hitchcock, Christopher, 2001, "The Intransitivity of Causation Revealed in Equations and Graphs", *Journal of Philosophy*, 98: 273-99.
- Hobart, R.E., 1934, "Free Will as Involving Determination and Inconceivable without It", *Mind*, 63: 1-27.
- Hofer, Carl, 2002, "Freedom from the Inside Out", in *Time, Reality, and Experience*, Craig Callender, ed., Cambridge: Cambridge University Press, 201-222.
- Holbach, Baron, 1770, "The Illusion of Free Will", from *The System of Nature*; page reference is to the reprint in *The Experience of Philosophy*, D. Kolak and R. Martin (eds.), Oxford: Oxford University Press, 2002, 176-181.
- Horgan, Terence, 1985, "Compatibilism and the Consequence Argument", *Philosophical Studies*, 47: 339-56.
- Horwich, Paul, 1987, *Asymmetries in Time*, Cambridge, MA: MIT Press.
- Huemer, Michael, 2000, "Van Inwagen's Consequence Argument", *The Philosophical Review*, 109: 525-544.
- James, William, 1884, "The Dilemma of Determinism", *An Address to the Harvard Divinity Students*; page reference is to the reprint in *The Will to Believe and Other Essays*, New York: Dover Publications, 1956, 145-183.
- Kane, Robert, 1989, "Two Kinds of Incompatibilism", in O'Connor 1995, 115-150.
- —, 1996, *The Significance of Free Will*, Oxford: Oxford University Press.

- —, 1999, “Responsibility, Luck, and Chance: Reflections on Free Will and Indeterminism”, *Journal of Philosophy*, 96: 217-240. Reprinted in Watson 2003.
- Kane, Robert, 2005, *A Contemporary Introduction to Free Will*, New York: Oxford University Press.
- Kane, Robert, (ed.), 2002A, *Free Will*, Oxford: Blackwell.
- —, 2002B, *The Oxford Handbook of Free Will*, Oxford: Oxford University Press.
- Kapitan, Tomis, 1991, “Ability and Cognition: a Defense of Compatibilism”, *Philosophical Studies*, 63: 231-43.
- —, 1996, “Incompatibilism and ambiguity in the practical modalities”, *Analysis*, 56: 102-110.
- —, 2000, “Autonomy and Manipulated Freedom”, *Philosophical Perspectives* 14 (*Action and Freedom*), James Tomberlin (ed.), Oxford: Blackwell, 81-104.
- —, 2002, “A Master Argument for Incompatibilism?”, in Kane 2002B, 127-157.
- Lamb, James, 1977, “On a Proof of Incompatibilism”, *The Philosophical Review*, 86: 20-35.
- —, 1993, “Evaluative Compatibilism and the Principle of Alternative Possibilities”, *Journal of Philosophy*, 90: 497-516.
- Lehrer, Keith, 1968, “Cans without Ifs”, *Analysis*, 29: 29-32.
- —, 1980, “Preferences, Conditionals, and Freedom”, in *Metamind*, K. Lehrer (ed.), Oxford: Clarendon Press, 1990, 79-95.
- Lewis, David, 1973, *Counterfactuals*, Harvard: Harvard University Press.
- —, 1976, “The Paradoxes of Time Travel”, *American Philosophical Quarterly*, 13: 145-52
- —, 1979, “Counterfactual Dependence and Time's Arrow”, *Noûs*, 13: 455-76.
- —, 1981, “Are We Free to Break the Laws?”, *Theoria*, 47: 113-21. Reprinted in Watson 2003.
- —, 1986, “Against Overlap”, in *On the Plurality of Worlds*, Oxford: Blackwell, 198-209.
- Loewer, Barry, 1996a, “Humean Supervenience”, *Philosophical Topics*, 24: 101-27.
- —, 1996b, “Freedom from Physics: Quantum Mechanics and Free Will”, *Philosophical Topics*, 24: 91-112 .
- Mackie, Penelope, 2003, “Fatalism, Incompatibilism, and the Power to Do Otherwise”, *Noûs*, 37: 672-689.
- Markosian, Ned, 1999, “A Compatibilist Version of the Theory of Agent Causation”, *Pacific Philosophical Quarterly*, 80: 257-277.
- McKay, Thomas and Johnson, David, 1996, “A Reconsideration of an Argument against Compatibilism”, *Philosophical Topics*, 24: 113-22.
- McKenna, Michael, 2001, “Source Incompatibilism, Ultimacy, and the Transfer of Non-Responsibility”, *American Philosophical Quarterly*, 38: 37-52.
- McKenna, Michael, 2004, “Responsibility and Globally Manipulated Agents”, *Philosophical Topics*, 32: 169-192, John Fischer (ed.).
- —, 2008, “A Hard-line Reply to Pereboom's Four-Case Manipulation Argument”, *Philosophy and Phenomenological Research*, 76.3.
- Mele, Alfred, 1995, *Autonomous Agents*, New York: Oxford University Press.
- —, 2006, *Free Will and Luck*, New York: Oxford University Press.
- Moore, G.E. , 1912, “Free Will” in *Ethics*, Oxford: Oxford University Press, Chapter 6.

- Nagel, Thomas, 1979, "Moral Luck", in his *Mortal Questions*, New York: Cambridge University Press, 24-38.
- Narveson, Jan, 1977, "Compatibilism Defended", *Philosophical Studies*, 32: 83-88.
- Nelkin, Dana, 2001, "The Consequence Argument and the Mind argument", *Analysis*, 61: 107-115.
- Nelkin, Dana and Rickless, Samuel, 2002, "Warfield's New Argument for Incompatibilism", *Analysis*, 62: 104-107.
- O'Connor, Timothy, 1993, "On the Transfer of Necessity", *Noûs*, 27: 204-218.
- —, 1995, "Agent Causation", in O'Connor 1995, 173-200. Reprinted in Watson 2003.
- —, 1995, ed., *Agents, Causes, and Events: Essays on Indeterminism and Free Will*, New York: Oxford University Press.
- —, 2000, *Persons and Causes: the Metaphysics of Free Will*, New York: Oxford University Press.
- Pereboom, Derk, 1995, "Determinism al Dente", *Noûs*, 29: 21-45.
- —, 2001, *Living Without Free Will*, Cambridge: Cambridge University Press.
- —, 2005, "Defending Hard Incompatibilism", *Midwest Studies in Philosophy*, 29: 228-247.
- Perry, John, 2004, "Compatibilist Options", in *Freedom and Determinism*, J. Campbell, M. O'Rourke, and D. Shier (eds.), MIT Press: Bradford Books, 231-254.
- Rosen, Gideon, 2002, "The Case for Incompatibilism", *Philosophy and Phenomenological Research*, 64: 699-706.
- Schaffer, Jonathan, 2008, "Causation and Laws of Nature: Reductionism", in *Contemporary Debates in Metaphysics*, J. Hawthorne, T. Sider, D. Zimmerman (eds.), Malden, MA: Blackwell Publishing, 82-107.
- Schneider, Benjamin, 2004, "Compatibilism and the Notion of Rendering Something False" *Philosophical Studies*, 117: 409-428.
- Shoemaker, Sydney, 1984, "Causality and Properties" in his *Identity, Cause, and Mind: Philosophical Essays*, London: Cambridge University Press, 206-233.
- Sider, Theodore, 2001, *Four-Dimensionalism*, Oxford: Clarendon Press.
- Slote, Michael, 1982, "Selective Necessity and the Free Will Problem", *Journal of Philosophy*, 79: 5-24.
- Sobel, Howard, 1998, *Puzzles for the Will*, Toronto: University of Toronto Press.
- Smilansky, Saul, 1993, "Does the Free Will Debate Rest on a Mistake?", *Philosophical Papers*, 22: 173-187.
- —, 2000, *Free Will and Illusion*, Oxford: Clarendon Press.
- Smith, Michael, 1997, "A Theory of Freedom and Responsibility", in *Ethics and Practical Reason*, G. Cullity (ed.), New York, Clarendon Press, 293-319.
- —, 2004, "Rational Capacities", in his *Ethics and the A Priori: Selected Essays on Moral Psychology and Meta-Ethics*, New York: Cambridge University Press, 114-135.
- Stone, Jim, 1998, "Free Will as a Gift from God: A New Compatibilism", *Philosophical Studies*, 92: 257-281.
- Strawson, Galen, 1986, *Freedom and Belief*, Oxford: Clarendon Press.
- —, 1994, "The Impossibility of Moral Responsibility", *Philosophical Studies*, 75: 5-24. Reprinted in Watson 2003.
- —, 2002, "The Bounds of Freedom", in Kane 2002B, 441-460.

- Strawson, P.F., 1962, "Freedom and Resentment", *Proceedings of the British Academy*, 68: 187-211. Reprinted in Watson 2003.
- Swartz, Norman, 1986, *The Concept of Physical Law*, Cambridge: Cambridge University Press.
- Taylor, Richard, 1960, *Action and Purpose*, Englewood Cliffs, N.J.: Prentice-Hall.
- —, 1967, "Determinism", *The Encyclopedia of Philosophy*, edited by Paul Edwards, New York: MacMillan.
- —, 1992, *Metaphysics*, Englewood Cliffs, NJ: Prentice-Hall.
- Tomberlin, James, (ed.), 2000, *Philosophical Perspectives*, 14 (*Action and Freedom*), Malden, MA: Blackwell Publishing.
- van Inwagen, Peter, 1975, "The Incompatibility of Free Will and Determinism", *Philosophical Studies*, 25: 185-99. Reprinted in Pereboom 1997 and Kane 2002.
- —, 1983, *An Essay on Free Will*, Oxford: Clarendon Press.
- —, 1993, *Metaphysics*, Boulder, CO: Westview Press.
- —, 1998, "The Mystery of Metaphysical Freedom", in *Metaphysics: The Big Questions*, P. van Inwagen and D. Zimmerman (eds.), Malden, MA: Blackwell Publishing, 365-374.
- —, 2000, "Free Will Remains a Mystery", in Tomberlin 2000, 1-19. Reprinted in Kane 2002B.
- —, 2004A, "Freedom to Break the Laws", in *Midwest Studies in Philosophy*, 28: 334-350.
- —, 2004B, "Van Inwagen on Free Will ", in *Freedom and Determinism*, ed. J. Campbell, M. O'Rourke, and D. Shier, MIT Press: Bradford Books, 213-230.
- Vargas, Manuel, 2004, "Libertarianism and Skepticism about Free Will: Some Arguments against Both", *Philosophical Topics*, 32: 403-426.
- Vihvelin, Kadri, 1990, "Freedom, Necessity, and Laws of Nature as Relations between Universals", *Australasian Journal of Philosophy*, 68: 371-381.
- —, 2000a, "Libertarian Compatibilism", in Tomberlin 2000, 139-166.
- —, 2000b, "Freedom, Foreknowledge, and the Principle of Alternate Possibilities", *Canadian Journal of Philosophy*, 30: 1-24.
- —, 2004, "Free Will Demystified: A Dispositional Account", *Philosophical Topics*, 32: 427-450 (edited by John M. Fischer).
- —, 2008, "Compatibilism, Incompatibilism, and Impossibilism", in *Contemporary Debates in Metaphysics*, John Hawthorne, Theodore Sider, and Dean Zimmerman (eds.), Malden, MA: Blackwell Publishing, 303-318.
- Wallace, R. Jay, 1994, *Responsibility and the Moral Sentiments*, Harvard: Harvard University Press.
- —, 2002, "Replies", *Philosophy and Phenomenological Research*, 64: 707-727.
- Warfield, Ted. A., 2000, "Causal Determinism and Human Freedom are Incompatible: A New Argument for Incompatibilism", in Tomberlin 2000, 167-180.
- —, 2003, "Compatibilism and Incompatibilism: Some Arguments", in *The Oxford Handbook of Metaphysics*, Michael J. Loux and Dean W. Zimmerman (eds.), Oxford: Oxford University Press, 613-630.
- Watson, Gary, 1975, "Free Agency", *Journal of Philosophy*, 72: 205-220. Reprinted in Watson 2003 and 2004.
- —, 1987, "Free Action and Free Will", *Mind*, 96: 145-72. Reprinted in Watson 2004.

- —, 1995, “Free Will”, in *A Companion to Metaphysics*, J. Kim and E. Sosa (eds.), Oxford: Blackwell, 175-182.
- —, 2004, *Agency and Accountability: Selected Essays*, Oxford: Clarendon Press.
- Watson, Gary, (ed.), 2003, *Free Will*, Oxford: Oxford University Press.
- Wegner, Daniel, 2003, *The Illusion of Conscious Will*, MIT Press: Bradford Books.
- Westphal, Jonathan, 2003, “A New Way with the Consequence Argument”, *Analysis*, 63: 209-212.
- —, 2006, “The future and truth-value links: a common sense view”, *Analysis*, 66: 1-9.
- Wolf, Susan, 1990, *Freedom within Reason*, Oxford: Oxford University Press.
- Widerker, David, 1987, “On an Argument for Incompatibilism”, *Analysis*, 47: 37-41.
- Widerker, David and McKenna, Michael (eds.), 2003, *Moral Responsibility and Alternative Possibilities*, Burlington, VT: Ashgate Publishing Company.
- Wiggins, David, 1973, “Towards a Reasonable Libertarianism”, in *Essays on Freedom and Action*, Ted Honderich (ed.), London: Routledge & Kegan Paul, 31-62. Reprinted in Watson 2003.
- Zimmerman, Michael J., 1988, *An Essay on Moral Responsibility*, Totowa, N.J.: Rowman & Littlefield.

## Other Internet Resources

- [The Determinism and Freedom Philosophy Website](#), edited by Ted Honderich (University College London).
- [Bibliography on Free will](#), maintained by David Chalmers (ANU).

## Related Entries

[action](#) | [causation: counterfactual theories of](#) | [causation: the metaphysics of](#) | [determinism: causal](#) | [fatalism](#) | [free will](#) | [incompatibilism: \(nondeterministic\) theories of free will](#) | [moral responsibility](#) | [providence, divine](#) | [time](#)

Copyright © 2007 by  
Kadri Vihvelin <[vihvelin@rcf.usc.edu](mailto:vihvelin@rcf.usc.edu)>