

# Internet Encyclopedia of Philosophy

A Peer-Reviewed Academic Resource



A  
B  
C  
D  
E  
F  
G  
H  
I  
J  
K  
L  
M  
N  
O  
P  
Q  
R  
S  
T  
U  
V  
W  
X  
Y  
Z

## Free Will

Most of us are certain that we have free will, though what exactly this amounts to is much less certain. According to [David Hume](#), the question of the nature of free will is “the most contentious question of metaphysics.” If this is correct, then figuring out *what free will is* will be no small task indeed. Minimally, to say that an agent has free will is to say that the agent has the capacity to choose his or her course of action. But animals seem to satisfy this criterion, and we typically think that only persons, and not animals, have free will. Let us then understand free will as the capacity unique to persons that allows them to control their actions. It is controversial whether this minimal understanding of what it means to have a free will actually requires an agent to have a specific faculty of will, whether the term “free will” is simply shorthand for other features of persons, and whether there really is such a thing as free will at all.

This article considers why we should care about free will and how freedom of will relates to freedom of action. It canvasses a number of the dominant accounts of what the will is, and then explores the persistent question of the relationship between free will and causal determinism, articulating a number of different positions one might take on the issue. For example, does determinism imply that there is no free will, as the incompatibilists argue, or does it allow for free will, as the compatibilists argue? This article explores several influential arguments that have been given in favor of these two dominant positions on the relationship between free will and causal determinism. Finally, there is a brief examination of how free will relates to theological determinism and logical determinism.

### Table of Contents

- [Free Will, Free Action and Moral Responsibility](#)
- [Accounts of the Will](#)
  - [Faculties Model of the Will](#)
  - [Hierarchical Model of the Will](#)
  - [Reasons-Responsive View of the Will](#)
- [Free Will and Determinism](#)
  - [The Thesis of Causal Determinism](#)
  - [Determinism, Science and “Near Determinism”](#)
  - [Compatibilism, Incompatibilism, and Pessimism](#)
- [Arguments for Incompatibilism \(or Arguments against Compatibilism\)](#)
  - [The Consequence Argument](#)
  - [The Origination Argument](#)
  - [The Relation between the Arguments](#)
- [Arguments for Compatibilism \(or Arguments against Incompatibilism\)](#)
  - [Rejecting the Incompatibilist Arguments](#)
  - [Frankfurt’s Argument against “the Ability to Do Otherwise”](#)
  - [Strawson’s Reactive Attitudes](#)
- [Related Issues](#)
  - [Theological Determinism](#)
  - [Logical Determinism](#)
- [References and Further Reading](#)

# 1. Free Will, Free Action and Moral Responsibility

Why should we even care whether or not agents have free will? Probably the best reason for caring is that free will is closely related to two other important philosophical issues: freedom of action and moral responsibility. However, despite the close connection between these concepts, it is important not to conflate them.

We most often think that an agent's free actions are those actions that she does as a result of exercising her free will. Consider a woman, Allison, who is contemplating a paradigmatic free action, such as whether or not to walk her dog. Allison might say to herself, "I know I should walk the dog—he needs the exercise. And while I don't really want to walk him since it is cold outside, I think overall the best decision to make is that I should take him for a walk." Thus, we see that one reason we care about free will is that it seems necessary for free action—Allison must first decide, or choose, to walk the dog before she actually takes him outside for his walk. If we assume that human actions are those actions that result from the rational capacities of humans, we then see that the possibility of free action depends on the possibility of free will: to say that an agent acted freely is minimally to say that the agent was successful in carrying out a free volition or choice.

Various philosophers have offered just such an account of freedom. [Thomas Hobbes](#) suggested that freedom consists in there being no external impediments to an agent doing what he wants to do: "A *free agent* is he that *can do as he will*, and *forbear* as he will, and that *liberty is the absence of external impediments*." In *An Enquiry Concerning Human Understanding*, David Hume thought that free will (or "liberty," to use his term) is simply the "power of acting or of not acting, according to the determination of the will: that is, if we choose to remain at rest, we may; if we choose to move, we also may.... This hypothetical liberty is universally allowed to belong to everyone who is not a prisoner and in chains." This suggests that freedom is simply the ability to select a course of action, and an agent is free if he is not being prevented by some external obstacle from completing that course of action. Thus, Hobbes and Hume would hold that Allison is free to walk her dog so long as nothing prevents her from carrying out her decision to walk her dog, and she is free not to walk her dog so long as nothing would compel her to walk her dog if she would decide not to.

However, one might still believe this approach fails to make an important distinction between these two related, but conceptually distinct, kinds of freedom: freedom of will versus freedom of action. This distinction is motivated by the apparent fact that agents can possess free will without also having freedom of action. Suppose that before Allison made the choice to walk the dog, she was taking a nap. And while Allison slept, there was a blizzard that moved through the area. The wind has drifted the snow up against the front of her house so that it is impossible for Allison to get out her front door and walk her dog even if she wanted to. So here we have a case involving free will, because Allison has chosen to take the dog for a walk, but not involving free action, because Allison is not able to take her dog for a walk.

Whether or not one can have freedom of action without free will depends on one's view of what free will is. Also, the truth of causal determinism would not entail that agents lack the freedom to do what they want to do. An agent could do what she wants to do, even if she is causally determined to do that action. Thus, both Hobbes and Hume are rightly characterized as compatibilists.

Even if there is a distinction between freedom of will and freedom of action, it appears that free will is necessary for the performance of free actions. If Allison is brainwashed during her nap to want to walk her dog, then even if no external impediment prevents her from carrying through with this decision, we would say that her taking the dog for a walk is not a free action. Presumably, the reason why it would not be a free action is because, in the case of brainwashing, Allison's decision does not arise from her free will. Thus, it looks like free will might be a necessary condition for free action, even if the two are distinct. In what follows, the phrase "acting with free will" means engaging in an action as the result of the utilization of free will. Use of the phrase does not deny the distinction between free will and free action.

The second reason to care about free will is that it seems to be required for moral responsibility. While there are various accounts of what exactly moral responsibility is, it is widely agreed that moral responsibility is distinct from causal responsibility. Consider a falling branch that lands on a car, breaking its window. While the branch is causally responsible for the broken window, it is not morally responsible for it because branches are not moral agents. Depending on one's account of causation, it also might be possible to be morally responsible for an event or state of affairs even if one is not causally responsible for that same event or state of affairs. For present purposes, let us simply say that an agent is morally responsible for an event or state of affairs only if she is the appropriate recipient of moral praise or moral blame for that event or state of affairs (an agent can thus be morally responsible even if no one, including herself, actually does blame or praise her for her actions). According to the dominant view of the relationship between free will and moral responsibility, if an agent does not have free will, then that agent is not morally responsible for her actions. For example, if Allison is coerced into doing a morally bad act, such as stealing a car, we shouldn't hold her morally responsible for this action since it is not an action that she did of her own free will.

Some philosophers do not believe that free will is required for moral responsibility. According to John Martin Fischer, human agents do not have free will, but they are still morally responsible for their choices and actions. In a nutshell, Fischer thinks that the kind of control needed for moral responsibility is weaker than the kind of control needed for free will. Furthermore, he thinks that the truth of causal determinism would preclude the kind of control needed for free will, but that it wouldn't preclude the kind of control needed for moral responsibility. See Fischer (1994). As this example shows, virtually every issue pertaining to free will is contested by various philosophers.

However, many think that the significance of free will is not limited to its necessity for free action and moral responsibility. Various philosophers suggest that free will is also a requirement for agency, rationality, the autonomy and dignity of persons, creativity, cooperation, and the value of friendship and love [see Anglin (1990), Kane (1998) and Ekstrom (1999)]. We thus see that free will is central to many philosophical issues.

## 2. Accounts of the Will

Nearly every major figure in the history of philosophy has had something or other to say about free will. The present section considers three of the most prominent theories of what the will is.

### a. Faculties Model of the Will

The faculties model of the will has its origin in the writings of ancient philosophers such as [Plato](#) and [Aristotle](#), and it was the dominant view of the

will for much of medieval and modern philosophy [see Descartes (1998) and the discussion of Aquinas in Stump (2003)]. It still has numerous proponents in the contemporary literature. What is distinct about free agents, according to this model, is their possession of certain powers or capacities. All living things possess some capacities, such as the capacities for growth and reproduction. What is unique about free agents, however, is that they also possess the capacities for intellection and volition. Another way of saying this is that free agents alone have the faculties of intellect and will. It is in virtue of having these additional faculties, and the interaction between them, that agents have free will.

The intellect, or the rational faculty, is the power of cognition. As a result of its cognitions, the intellect presents various things to the will as good under some description. To return to the case of Allison contemplating walking her dog, Allison's intellect might evaluate walking the dog as good for the health of the dog. Furthermore, all agents that have an intellect also have a will. The will, or the volitional faculty, is an appetite for the good; that is, it is naturally drawn to goodness. The will, therefore, cannot pursue an option that the intellect presents as good in no way. The will is also able to command the other faculties; the will can command the body to move or the intellect to consider something. In the case of Allison, the will could command the body to pick up the leash, attach it to the dog, and go outside for a walk. As Aquinas, a proponent of this view of the will, puts it: "Only an agent endowed with an intellect can act with a judgment which is free, in so far as it apprehends the common note of goodness; from which it can judge this or the other thing to be good. Consequently, wherever there is intellect, there is free will" (*Summa Theologiae*, q. 59 a. 3). Thus, through the interaction between the intellect and will, an agent has free will to pursue something that it perceives as good.

## b. Hierarchical Model of the Will

A widely influential contemporary account of the will is Harry Frankfurt's hierarchical view of the will [see Frankfurt (1971)]. This account is also sometimes called a "structuralist" or "mesh" account of the will, since a will is free if it has a certain internal structure or "mesh" among the various levels of desires and volitions. According to the hierarchical model, agents can have different kinds of desires. Some desires are desires to do a particular action; for example, Allison may desire to go jogging. Call these desires "1st order desires." But even if Allison doesn't desire to go jogging, she may nevertheless desire to be the kind of person who desires to go jogging. In other words, she may desire to have a certain 1st order desire. Call desires of this sort "2nd order desires." If agents also have further desires to have particular 2nd order desires, one could construct a seemingly infinite hierarchy of desires.

Not all of an agent's desires result in action. In fact, if one has conflicting desires, then it is impossible for an agent to satisfy all her desires. Suppose that Allison not only desires to run, but that she also desires to stay curled up in bed, where it is nice and warm. In such a case, Allison cannot fulfill both of her 1st order desires. If Allison decides to act on her desire to run, we say that her desire to run has moved her to action. An effective desire of this sort is called a volition; a volition is a desire that moves the agent all the way to action. Similarly, one can differentiate between a mere 2nd order desire (simply a desire to have a certain desire) and a 2nd order volition (a desire for a desire to become one's will, or a desire for a desire to become a volition). According to the hierarchical view of the will, free will consists in having 2nd order volitions. In other words, an agent has a free will if she is able to have the sort of will that she wants to have. An agent acts on her own free will if her action is the result of a 1st order desire that she wants to become a 1st order volition.

Hierarchical views of the will are problematic, however, because it looks as if certain sorts of questionable manipulation can be compatible with this view's account of free will. According to the view under consideration, Allison has free will with regard to going jogging if she has a 2nd order desire that her 1st order desire to go jogging will move her to go jogging. Nothing in this account, however, depends on how she got these desires. Even if she were manipulated, via brainwashing, for example, into having her 2nd order desire for her 1st order desire to go running become her will, Allison has the right "mesh" between her various orders of desires to qualify as having free will. This is an untoward consequence. While more robust hierarchical accounts of the will have the resources for explaining why Allison might not be free in this case, it is widely agreed that cases of manipulation and coercion are problematic for solely structural accounts of the will [see Ekstrom (1999), Fischer (1994), Kane, (2005), Pereboom (2001) and van Inwagen (1983)].

## c. Reasons-Responsive View of the Will

A third treatment of free will takes as its starting point the claim that agency involves a sensitivity to certain reasons. An agent acts with free will if she is responsive to the appropriate rational considerations, and she does not act with a free will if she lacks such responsiveness. To see what such a view amounts to, consider again the case of Allison and her decision to walk her dog. A reasons-responsive view of the will says that Allison's volition to walk her dog is free if, had she had certain reasons for not walking her dog, she would not have decided to walk her dog. Imagine what would have happened had Allison turned on the television after waking from her nap and learned of the blizzard before deciding to walk her dog. Had she known of the blizzard, she would have had a good reason for deciding not to walk her dog. Even if such reasons never occur to her (that is, if she doesn't learn of the blizzard before her decision), her disposition to have such reasons influence her volitions shows that she is responsive to reasons. Thus, reasons-responsive views of the will are essentially dispositional in nature.

Coercion and manipulation undermine free will, on this view, in virtue of making agents not reasons-responsive. If Allison has been brainwashed to walk the dog at a certain time, then even if she were to turn on the news and sees that it is snowing, she would attempt to walk the dog despite having good reasons not to. Thus, manipulated agents are not reasons-responsive, and in virtue of this lack free will. [See Fischer and Ravizza (1998) for one of the primary reasons-responsive views of free will.]

# 3. Free Will and Determinism

## a. The Thesis of Causal Determinism

Most contemporary scholarship on free will focuses on whether or not it is compatible with causal determinism. Causal determinism is sometimes also called "nomological determinism." It is important to keep causal determinism distinct from other sorts of [determinism](#), such as logical determinism or theological determinism (to be discussed below). Causal determinism (hereafter, simply "determinism") is the thesis that the course of the future is entirely determined by the conjunction of the past and the laws of nature. Imagine a proposition that completely describes the way that the entire universe was at some point in the past, say 100 million years ago. Let us call this proposition "P." Also imagine a proposition that expresses the conjunction of all the laws of nature; call this proposition "L." Determinism then is the thesis that the conjunction of P and L entails a

unique future. Given P and L, there is only one possible future, one possible way for things to end up. To make the same point using possible world semantics, determinism is the thesis that all the states of affairs that obtain at some time in the past, when conjoined with the laws of nature, entail which possible world is the actual world. Since a possible world includes those states of affairs that will obtain, the truth of determinism amounts to the thesis that the past and the laws of nature entail what states of affairs will obtain in the future, and that only those states of affairs entailed by the past and the laws will in fact obtain.

A system's being determined is different from its being predictable. It is possible for determinism to be true and for no one to be able to predict the future. The fact that no human agent knows or is able to know future truths has no bearing on whether there are future truths entailed by the conjunction of the past and the laws. However, there is a weaker connection between the thesis of determinism and the predictability of the future. If determinism were true, then a being with a complete knowledge of P and L and with sufficient intellectual capacities should be able to infallibly predict the way that the future will turn out. However, given that we humans lack both the relevant knowledge and the intellectual capacities required, the fact that we are not able to predict the future is not evidence for the falsity of determinism.

## b. Determinism, Science and “Near Determinism”

Most philosophers agree that whether or not determinism is true is a contingent matter; that is, determinism is neither necessarily true nor necessarily false. If this is so, then whether or not determinism is true becomes an empirical matter, to be discovered by investigating the way the world is, not through philosophical argumentation. This is not to deny that the truth of determinism would have metaphysical implications. For one, the truth of determinism would entail that the laws of nature are not merely probabilistic—for if they were, then the conjunction of the past and the laws would not entail a unique future. Furthermore, as we shall see shortly, philosophers care very much about what implications the truth of determinism would have for free will. But the point to note is that if the truth of determinism is a contingent truth about the way the world actually is, then scientific investigation should give us insight into this matter. Let us say that a possible world is deterministic if causal determinism is true in that world. There are two ways that worlds could fail to be deterministic. As already noted, if the laws of nature in a given world were probabilistic, then such a world would not be deterministic. Secondly, if there are entities within a world that are not fully governed by the laws of nature, then even if those laws are themselves deterministic, that world would not be deterministic.

Some scientists suggest that certain parts of physics give us reason to doubt the truth of determinism. For example, the standard interpretation of Quantum Theory, the Copenhagen Interpretation, holds that the laws governing nature are indeterministic and probabilistic. According to this interpretation, whether or not a small particle such as a quark swerves in a particular direction at a particular time is described properly only by probabilistic equations. Although the equations may predict the likelihood that a quark swerves to the left at a certain time, whether or not it actually swerves is indeterministic or random.

There are also deterministic interpretations of Quantum Theory, such as the Many-Worlds Interpretation. Fortunately, the outcome of the debate regarding whether Quantum Theory is most properly interpreted deterministically or indeterministically, can be largely avoided for our current purposes. Even if (systems of) micro-particles such as quarks are indeterministic, it might be that (systems involving) larger physical objects such as cars, dogs, and people are deterministic. It is possible that the only indeterminism is on the scale of micro-particles and that macro-objects themselves obey deterministic laws. If this is the case, then causal determinism as defined above is, strictly speaking, false, but it is “nearly” true. That is, we could replace determinism with “near determinism,” the thesis that despite quantum indeterminacy, the behaviors of all large physical objects—including all our actions—obey deterministic laws [see Honderich (2002), particularly chapter 6].

What would be the implications of the truth of either determinism or near determinism? More specifically, what would be the implications for questions of free will? One way to think about the implications would be by asking the following question: could we still be free even if scientists were to discover that causal determinism (or near determinism) is true?

## c. Compatibilism, Incompatibilism, and Pessimism

The question at the end of the preceding section (Could we have free will even if determinism is true?) is a helpful way to differentiate the main positions regarding free will. Compatibilists answer this question in the affirmative. They believe that agents could have free will even if causal determinism is true (or even if near determinism is true. In what follows, I will omit this qualification). In other words, the existence of free will in a possible world is compatible with that world being deterministic. For this reason, this position is known as “compatibilism,” and its proponents are called “compatibilists.” According to the compatibilist, it is possible for an agent to be determined in all her choices and actions and still make some of her choices freely.

According to “incompatibilists,” the existence of free will is incompatible with the truth of determinism. If a given possible world is deterministic, then no agent in that world has free will for that very reason. Furthermore, if one assumes that having free will is a necessary condition for being morally responsible for one's actions, then the incompatibility of free will and determinism would entail the incompatibility of moral responsibility and causal determinism.

There are at least two kinds of incompatibilists. Some incompatibilists think that determinism is true of the actual world, and thus no agent in the actual world possesses free will. Such incompatibilists are often called “hard determinists” [see Pereboom (2001) for a defense of hard determinism]. Other incompatibilists think that the actual world is not deterministic and that at least some of the agents in the actual world have free will. These incompatibilists are referred to as “libertarians” [see Kane (2005), particularly chapters 3 and 4]. However, these two positions are not exhaustive. It is possible that one is an incompatibilist, thinks that the actual world is not deterministic, and yet still thinks that agents in the actual world do not have free will. While it is less clear what to call such a position (perhaps “free will deniers”), it illustrates that hard determinism and libertarianism do not exhaust the ways to be an incompatibilist. Since all incompatibilists, whatever their stripe, agree that the falsity of determinism is a necessary condition for free will, and since compatibilists deny this assertion, the following sections speak simply of incompatibilists and compatibilists.

It is also important to keep in mind that both compatibilism and incompatibilism are claims about possibility. According to the compatibilist, it is possible that an agent is both fully determined and yet free. The incompatibilist, on the other hand, maintains that such a state of affairs is impossible. But neither position by itself is making a claim about whether or not agents actually do possess free will. Assume for the moment that incompatibilism is true. If the truth of determinism is a contingent matter, then whether or not agents are morally responsible will depend on

whether or not the actual world is deterministic. Furthermore, even if the actual world is indeterministic, it doesn't immediately follow that the indeterminism present is of the sort required for free will (we will return to a similar point below when considering an objection to incompatibilism). Likewise, assume both that compatibilism is true and that causal determinism is true in the actual world. It does not follow from this that agents in the actual world actually possess free will.

Finally, there are free will pessimists [see Broad (1952) and G. Strawson (1994)]. Pessimists agree with the incompatibilists that free will is not possible if determinism is true. However, unlike the incompatibilists, pessimists do not think that indeterminism helps. In fact, they claim, rather than helping support free will, indeterminism undermines it. Consider Allison contemplating taking her dog for a walk. According to the pessimist, if Allison is determined, she cannot be free. But if determinism is false, then there will be indeterminacy at some point prior to her action. Exactly where one locates this indeterminacy will depend on one's particular view of the nature of free will. Let us assume that that indeterminacy is located in which reasons occur to Allison. It is hard to see, the pessimist argues, how this indeterminacy could enhance Allison's free will, for it the occurrence of her reasons is indeterministic, then having those reasons is not within Allison's control. But if Allison decides on the basis of whatever reasons she does have, then her volition is based upon something outside of her control. It is based instead on chance. Thus, pessimists think that the addition of indeterminism actually makes agents lack the kind of control needed for free will. While pessimism might seem to be the same position as that advocated by free will deniers, pessimism is a stronger claim. Free will deniers think that while free will is possible, it just isn't actual: agents in fact don't have free will. Pessimists, however, have a stronger position, thinking that free will is impossible. Not only do agents lack free will, there is no way that they could have it [see G. Strawson (1994)]. The only way to preserve moral responsibility, for the pessimist, is thus to deny that free will is a necessary condition for moral responsibility.

As pessimism shows us, even a resolution to the debate between compatibilists and incompatibilists will not by itself solve the debate about whether or not we actually have free will. Nevertheless, it is to this debate that we now turn.

## 4. Arguments for Incompatibilism (or Arguments against Compatibilism)

Incompatibilists say that free will is incompatible with the truth of determinism. Not all arguments for incompatibilism can be considered here; let us focus on two major varieties. The first variety is built around the idea that having free will is a matter of having a choice about certain of our actions, and that having a choice is a matter of having genuine options or alternatives about what one does. The second variety of arguments is built around the idea that the truth of determinism would mean that we don't cause our actions in the right kind of way. The truth of determinism would mean that we don't originate our actions in a significant way and our actions are not ultimately controlled by us. In other words, we lack the ability for self-determination. Let us consider a representative argument from each set.

### a. The Consequence Argument

The most well-known and influential argument for incompatibilism from the first set of arguments is called the "Consequence Argument," and it has been championed by Carl Ginet and Peter van Inwagen [see Ginet (1966) and van Inwagen (1983)]. The Consequence Argument is based on a fundamental distinction between the past and the future. First, consider an informal presentation of this argument. There seems to be a profound asymmetry between the past and the future based on the direction of the flow of time and the normal direction of causation. The future is open in a way that the past is not. It looks as though there is nothing that Allison can now do about the fact that Booth killed Lincoln, given that Lincoln was assassinated by Booth in 1865.

This point stands even if we admit the possibility of time travel. For if time travel is possible, Allison can influence what the past became, but she cannot literally change the past. Consider the following argument:

The proposition "Lincoln was assassinated in 1865" is true.  
 If Allison travels to the past, she could prevent Lincoln from being assassinated in 1865 (temporarily assumed for *reductio* purposes).  
 If Allison were to travel to the past and prevent Lincoln from being assassinated in 1865, the proposition "Lincoln was assassinated in 1865" would be false.  
 A proposition cannot both be true and false.  
 Therefore, 2 is false.

So, at most the possibility of time travel allows for agents to have causal impact on the past, not for agents to change what has already become the past. The past thus appears to be fixed and unalterable. However, it seems that the same is not true of the future, for Allison can have an influence on the future through her volitions and subsequent actions. For example, if she were to invent a time machine, then she could, at some point in the future, get in her time machine and travel to the past and try to prevent Lincoln from being assassinated. However, given that he was assassinated, we can infer that her attempts would all fail. On the other hand, she could refrain from using her time machine in this way.

The asymmetry between past and future is illustrated by the fact that we don't deliberate about the past in the same way that we deliberate about the future. While Allison might deliberate about whether a past action was really the best action that she could have done, she deliberates about the future in a different way. Allison can question whether her past actions were in fact the best, but she can both question what future acts would be best as well as which future acts she should perform. Thus, it looks like the future is open to Allison, or up to her, in a way that the past is not. In other words, when an agent like Allison is using her free will, what she is doing is selecting from a range of different options for the future, each of which is possible given the past and the laws of nature. For this reason, this view of free will is often called the "Garden of Forking Paths Model."

The Consequence Argument builds upon this view of the fixed nature of the past to argue that if determinism is true, the future is not open in the way that the above reflections suggest. For if determinism is true, the future is as fixed as is the past. Remember from the above definition that determinism is the thesis the past (*P*) and the laws of nature (*L*) entail a unique future. Let "*F*" refer to any true proposition about the future. The Consequence argument depends on two modal operators, and two inference rules. Let the modal operator " $\Box$ " abbreviate "It is logically necessary that..." so that, when it operates on some proposition *p*, " $\Box p$ " abbreviates "It is logically necessary that *p*." Let the modal operator "*N*" be such that "*Np*" stands for "*p* is true and no one has, or ever had, any choice about whether *p* was true." Call the following two inference rules "Alpha" and "Beta:"

Alpha:  $\Box p$  implies *Np* Beta:  $\{Np \text{ and } N(p \rightarrow q)\}$  implies *Nq*

According to Alpha, if  $p$  is a necessary truth, then no one has, or ever had, any choice about whether  $p$  was true. Similarly, according to Beta, if no one has, or ever had, any choice about  $p$  being true, and no one has, or ever had, any choice that  $p$  entails  $q$ , then no one has, or ever had, any choice about whether  $q$  is true. To see the plausibility of Beta, consider the following application. Let  $p$  be the proposition “The earth was struck by a meteor weighing 100 metric tons one billion years ago,” and let  $q$  be the proposition “If the earth was struck by a meteor weighing 100 metric tons one billion years ago, then thousands of species went extinct.” Since I have no choice about such a meteor hitting in the past, and have no choice that if such meteor hits, it will cause thousands of species to go extinct, I have no choice that thousands of species went extinct. Beta thus looks extremely plausible. But if Beta is true, then we can construct an argument to show that if determinism is true, then I have no choice about anything, including my supposed free actions in the future. The argument begins with the definition of determinism given above:

$$(1) \Box\{(P \text{ and } L) \rightarrow F\}$$

Using a valid logical rule of inference (exportation), we can transform 1 into 2:

$$(2) \Box\{P \rightarrow (L \rightarrow F)\}$$

Applying Alpha, we can derive 3:

$$(3) N\{P \rightarrow (L \rightarrow F)\}$$

The second premise in the Consequence Argument is called the “fixity of the past.” No one has, or ever had, a choice about the true description  $P$  of the universe at some point in the distant past:

$$(4) NP$$

From 3, 4 and Beta, we can deduce 5:

$$(5) N(L \rightarrow F)$$

The final premise in the argument is the fixity of the laws of nature. No one has, or ever had, a choice about what the laws of nature are (try as I might, I cannot make the law of universal gravitation not be a law of nature):

$$(6) NL$$

And from 5 and 6, again using Beta, we can infer that no one has, or ever had, a choice about  $F$ :

$$(7) NF$$

Given that  $F$  was any true proposition about the future, the Consequence Argument concludes that if determinism is true, then no one has or ever had a choice about any aspect of the future, including what we normally take to be our free actions. Thus, if determinism is true, we do not have free will.

## b. The Origination Argument

The second general set of arguments for the incompatibility of free will and determinism builds on the importance of the source of a volition for free will. Again, it will be helpful to begin with an informal presentation of the argument before considering a formal presentation of it. According to this line of thought, an agent has free will when her volitions issue from the agent herself in a particular sort of way (say, her beliefs and desires). What is important for free will, proponents of this argument claim, is not simply that the causal chain for an agent’s volition goes through the agent, but that it *originates* with the agent. In other words, an agent acts with free will only if she originates her action, or if she is the ultimate source or first cause of her action [see Kane (1998)].

Consider again the claim that free will is a necessary condition for moral responsibility. What reflection on cases of coercion and manipulation suggests to us is that even if a coerced or manipulated agent is acting on her beliefs and desires, this isn’t enough for moral responsibility. We normally assume that coercion and certain forms of manipulation undercut an agent’s moral responsibility precisely because a coerced or manipulated agent isn’t the originator of her coerced action. If Allison is coerced into walking her dog via brainwashing, then her walking of the dog originates in the brainwashing, and not in Allison herself. Consider, then, the similarities between cases of coercion and manipulation, on the one hand, and the implications of the truth of determinism on the other. If determinism were true, it might be true that Allison chooses to walk her dog because of her beliefs and desires, but those beliefs and desires would themselves be the inevitable products of causal chains that began millions of years ago. Thus, a determined agent is at most a source, but not the ultimate source, of her volitions. According to proponents of this sort of argument for incompatibilism, the truth of determinism would mean that agents don’t cause their actions in the kind of way needed for free will and, ultimately, moral responsibility.

We can represent a formal version of the argument, called the “Origination Argument,” as follows:

An agent acts with free will only if she is the originator (or ultimate source) of her actions.

If determinism is true, then everything any agent does is ultimately caused by events and circumstances outside her control.

If everything an agent does is ultimately caused by events and circumstances beyond her control, then the agent is not the originator (or ultimate source) of her actions.

Therefore, if determinism is true, then no agent is the originator (or ultimate source) of her actions.

Therefore, if determinism is true, no agent has free will.

The Origination Argument is valid. So, in evaluating its soundness, we must evaluate the truth of its three premises. Premise 3 is clearly true, since for an agent to be an originator just is for that agent not to be ultimately determined by anything outside of herself. Premise 2 of this argument is

true by the definition of determinism. To reject the conclusion of the argument, one must therefore reject premise 1.

Earlier we briefly noted one account of free will which implicitly denies premise 1, namely the hierarchical model of free will. According to this model, an agent acts with free will so long as the causal chain for that action goes through the agent's 1st- and 2nd-order desires. One way of emphasizing the need for origination over-against such a hierarchical model is to embrace agent-causation. If premise 1 is true, then the agent's volition cannot be the product of a deterministic causal chain extended beyond the agent. What other options are there? Two options are that volitions are uncaused, or only caused indeterministically. It is difficult to see how an agent could be the originator or ultimate source of volitions if volitions are uncaused. Similarly, for reasons we saw above when discussing the free will pessimist, it looks as if indeterministic causation would undermine, rather than enhance, an agent's control over her volitions. For these reasons, some incompatibilists favor looking at the causation involved in volitions in a new light. Instead of holding that a volition is caused by a previous event (either deterministically or indeterministically), these incompatibilists favor saying that volitions are caused directly by agents. [For an extended defense of this view, see O'Connor, (2000).] They hold that there are two irreducibly different kinds of causation, event-causation and agent-causation, and the latter is involved in free will. Proponents of agent-causation propose that agents are enduring substances that directly possess the power to cause volitions. Although many philosophers question whether agent-causation is coherent, if it were coherent, then it would provide support for premise 1 of the Origination Argument.

### c. The Relation between the Arguments

The above way of delineating the Consequence and Origination Arguments may unfortunately suggest that the two kinds of arguments are more independent from each other than they really are. A number of incompatibilists have argued that agents originate their actions in the way required by premise 1 of the Origination Argument if and only if they have a choice about their actions in the way suggested by the Consequence Argument. In other words, if my future volitions are not the sort of thing that I have a choice about, then I do not originate those volitions. And as the above arguments contend, the truth of causal determinism threatens both our control over our actions and volitions, and our ability to originate those same actions and volitions. For if causal determinism is true, then the distant past, when joined with the laws of nature, is sufficient for every volition that an agent makes, and the causal chains that lead to those volitions would not begin within the agent. Thus, most incompatibilists think that having a choice and being a self-determiner go hand-in-hand. Robert Kane, for instance, argues that if agents have "ultimate responsibility" (his term for what is here called "origination" or "self-determination"), then they will also have alternative possibilities open to them. According to this line of argumentation, the power to cause one's own actions is not a distinct power from the power to choose and do otherwise. Thus, the two different kinds of arguments for incompatibilism may simply be two sides of the same coin [see Kane (1996) and (2005)].

## 5. Arguments for Compatibilism (or Arguments against Incompatibilism)

Having laid out representatives of the two most prominent arguments for incompatibilism, let's consider arguments in favor of compatibilism. In considering these kinds of arguments, it is pedagogically useful to approach them by using the arguments for incompatibilism. So, this section begins by considering ways that compatibilists have responded to the arguments given in the preceding section.

### a. Rejecting the Incompatibilist Arguments

As noted above, the Origination Argument for incompatibilism is valid, and two of its premises are above dispute. Thus, the only way for the compatibilist to reject the conclusion of the Origination Argument is to reject its first premise. In other words, given the definition of determinism, compatibilists must reject that free will requires an agent being the originator or ultimate source of her actions. But how might this be done? Most frequently, compatibilists motivate a rejection of the "ultimacy condition" of free will by appealing to either a hierarchical or reasons-responsive view of what the will is [see Frankfurt, (1971) and Fischer and Ravizza, (1998)]. If all that is required for free will, for example, is that a certain mesh between an agent's 1st-order volitions and 2nd-order desires, then such an account does not require that an agent be the originator of those desires. Furthermore, since the truth of determinism would not entail that agents don't have 1st and 2nd-order desires and volitions, a hierarchical account of the will is compatible with the truth of determinism. Similarly, if an agent has free will if she has the requisite level of reasons-responsiveness such that she would have willed differently had she had different reasons, ultimacy is again not required. Thus, if one adopts certain accounts of the will, one has reason for rejecting the central premise of the Origination Argument.

Compatibilists have a greater number of responses available to them with regard to the Consequence Argument. One way of understanding the N operator that figures in the Consequence Argument is in terms of having the ability to do otherwise. That is, to say that Allison has no choice about a particular action of hers is to say that she could not have performed a different action (or even no action at all). Incompatibilists can easily account for this ability to do otherwise. According to incompatibilists, an agent can be free only if determinism is false. Consider again the case of Allison. If determinism is false, even though Allison did choose to walk her dog, she could have done otherwise than walk her dog since the conjunction of *P* and *L* is not sufficient for her taking her dog for a walk. Compatibilists, however, can give their own account of the ability to do otherwise. For them, to say that Allison could have done otherwise is simply to say that Allison would have done otherwise had she willed or chosen to do so [see, for example, Chisholm (1967)]. Of course, if determinism is true, then the only way that Allison could have willed or chosen to do otherwise would be if either the past or the laws were different than they actually are. In other words, saying that an agent could have done otherwise is to say that the agent would have done otherwise in a different counterfactual condition. But saying this is entirely consistent with one way of understanding the ability to do otherwise. Thus, these compatibilists are saying that Allison has the ability to do something such that, had she done it, either the past or the laws of nature would have been different than they actually are. If *P* and *L* entail that the agent does some action *A*, then the agent's doing otherwise than *A* entails that either *P* or *L* would have been different than they actually are. Some compatibilists favor saying that agent's have this counterfactual power over the past, while others favor counterfactual power over the laws of nature [Compare Lewis (1981) and Fischer (1984)]. Regardless, adopting either strategy provides the compatibilist with a way of avoiding the conclusion of the Consequence Argument by denying either premise 4 or premise 6 of that argument. Furthermore, having such a power is not a hollow victory, for it demarcates a plausible difference between those actions an agent would have done even if she didn't want to (as in the case of coercion or manipulation) from those actions that an agent only would have done had she had certain beliefs and desires about that action. This view thus differentiates between those actions that were within the agent's power to bring about from those that were not.

A second compatibilist response to the Consequence Argument is to deny the validity of the inference rule Beta the argument uses. While there are several approaches to this, perhaps the most decisive is the following, called the principle of Agglomeration [see McKay and Johnson (1996)].

Using only the inference rules Alpha, Beta and basic rule of logical replacement, one can show that

(1)  $Np$

and

(2)  $Nq$

would entail

(3)  $N(p \text{ and } q)$

if Beta were valid. 1 and 2 do not entail 3, so Beta must be invalid.

To see why 3 does not follow from 1 and 2, consider the case of a coin-toss. If the coin-toss is truly random, then Allison has no choice regarding whether the coin (if flipped) lands heads. Similarly, she has no choice regarding whether the coin (again, if flipped) lands tails. For purposes of simplicity, let us stipulate that the coin cannot land on its side and, if flipped, must land either heads or tails. Let  $p$  above represent 'the coin doesn't land heads' and  $q$  represent 'the coin doesn't land tails'. If Beta were valid, then 1 and 2 would entail 3, and Allison would not have a choice about the conjunction of  $p$  and  $q$ ; that is, she wouldn't have a choice about the coin not landing heads and the coin not landing tails. If Allison didn't have a choice about the coin not landing heads and didn't have a choice about the coin not landing tails, then she wouldn't have a choice about the coin landing either heads or tails. But Allison does have a choice about this—after all, she can ensure that the coin lands either heads or tails by simply flipping the coin. So Allison does have a choice about the conjunction of  $p$  and  $q$ . Since Alpha and the relevant rules of logical replacement in the transformation from  $Np$  and  $Nq$  to  $N(p \text{ and } q)$  are beyond dispute, Beta must be invalid. Thus, the Consequent Argument for incompatibilism is invalid. [For an incompatibilist reply to the argument from Agglomeration, see Finch and Warfield (1998).]

### b. Frankfurt's Argument against "the Ability to Do Otherwise"

Two other arguments for compatibilism build on the freedom requirement for moral responsibility. If one can show that moral responsibility is compatible with the truth of determinism, and if free will is required for moral responsibility, one will have implicitly shown that free will is itself compatible with the truth of determinism. The first of these arguments for compatibilism rejects the understanding of having a choice as involving the ability to do otherwise mentioned above. While most philosophers have tended to accept that an agent can be morally responsible for doing an action only if she could have done otherwise, Harry Frankfurt has attempted to show that this requirement is in fact false. Frankfurt gives an example in which an agent does an action in circumstances that lead us to believe that the agent acted freely [Frankfurt (1969); for recent discussion, see Widerker and McKenna (2003)]. Yet, unbeknownst to the agent, the circumstances include some mechanism that would bring about the action if the agent did not perform it on her own. As it happens, though, the agent does perform the action freely and the mechanism is not involved in bringing about the action. It thus looks like the agent is morally responsible despite not being able to do otherwise. Here is one such scenario:

Allison is contemplating whether to walk her dog or not. Unbeknownst to Allison, her father, Lloyd, wants to insure that that she does decide to walk the dog. He has therefore implanted a computer chip in her head such that if she is about to decide not to walk the dog, the chip will activate and coerce her into deciding to take the dog for a walk. Given the presence of the chip, Allison is unable not to decide to walk her dog, and she lacks the ability to do otherwise. However, Allison does decide to walk the dog on her own.

In such a case, Frankfurt thinks that Allison is morally responsible for her decision since the presence of Lloyd and his computer chip play no causal role in her decision. Since she would have been morally responsible had Lloyd not been prepared to ensure that she decide to take her dog for a walk, why think that his mere presence renders her not morally responsible? Frankfurt concludes that Allison is morally responsible despite lacking the ability to do otherwise. If Frankfurt is right that such cases are possible, then even if the truth of determinism is incompatible with a kind of freedom that requires the ability to do otherwise, it is compatible with the kind of freedom required for moral responsibility.

### c. Strawson's Reactive Attitudes

In an influential article, Peter Strawson argues that the many of the traditional debates between compatibilists and incompatibilists (such as how to understand the ability to do otherwise) are misguided [P. Strawson (1963)]. Strawson thinks that we should instead focus on what he calls the reactive attitudes—those attitudes we have toward other people based on their attitudes toward and treatment of us. Strawson says that the hallmark of reactive attitudes is that they are "essentially natural human reactions to the good or ill will or indifference of others toward us, as displayed in *their* attitudes and actions." Examples of reactive attitudes include gratitude, resentment, forgiveness and love. Strawson thinks that these attitudes are crucial to the interpersonal interactions and that they provide the basis for holding individuals morally responsible. Strawson then argues for two claims. The first of these is that an agent's reactive attitudes would not be affected by a belief that determinism was true:

The human commitment to participation in ordinary interpersonal relationships is, I think, too thoroughgoing and deeply rooted for us to take seriously the thought that a general theoretical conviction might so change our world that, in it, there were no longer such things as inter-personal relationships as we normally understand them.... A sustained objectivity of inter-personal attitude, and the human isolation which that would entail, does not seem to be something of which human beings would be capable, even if some general truth were a theoretical ground for it.

Furthermore, Strawson also argues for a normative claim: the truth of determinism *should not* undermine our reactive attitudes. He thinks that there are two kinds of cases where it is appropriate to suspend our reactive attitudes. One involves agents, such as young children or the mentally disabled, who are *notmoral* agents. Strawson thinks that we should not have reactive attitudes toward non-moral agents. The second kind of case where it is appropriate to suspend our reactive attitudes are those in which while the agent is a moral agent, her action toward us is not connected to her agency in the correct way. For instance, while I might have the reactive attitude of resentment towards someone who bumps into me and makes me spill my drink, if I were to find out that the person was pushed into me, I would not be justified in resenting that individual. The truth of determinism, however, would neither entail that no agents are moral agents nor that none of an agent's actions are connected to her moral agency.



Thus, Strawson thinks, the truth of determinism should not undermine our reactive attitudes. Since moral responsibility is based on the reactive attitudes, Strawson thinks that moral responsibility is compatible with the truth of determinism. And if free will is a requirement for moral responsibility, Strawson's argument gives support to compatibilism.

## 6. Related Issues

The above discussion should help explain the perennial attraction philosophers have to the issues surrounding free will, particularly as it relates to causal determinism. However, free will is also intimately related to a number of other recurrent issues in the history of philosophy. In this final section, I will briefly articulate two other kinds of determinism and show how they are connected to free will.

### a. Theological Determinism

The debate about free will and causal determinism parallels, in many ways, another debate about free will, this one stemming from what is often called 'theological determinism'. Some religious traditions hold that God is ultimately responsible for everything that happens. According to these traditions, God's willing  $x$  is necessary and sufficient for  $x$ . But if He is ultimately responsible for everything in virtue of what He wills, then He is ultimately responsible for all the actions and volitions performed by agents. God's willing that Allison take the dog for a walk is thus necessary and sufficient for Allison taking the dog for a walk. But if this is true, it is hard to see how Allison could have free will. The problem becomes especially astute when considering tradition doctrines of eternal punishment. The traditional Christian doctrine of Hell, for example, is that Hell is a place of eternal punishment for non-repentant sinners. But if theological determinism is true, then whether or not agents repent is ultimately up to God, not to the agents themselves. This worry over free will thus gives rise to a particular version of the problem of evil: why does God not will that all come to faith, when His having such a will is sufficient for their salvation? [For a discussion of these, and related issues, see Helm, (1994).]

### b. Logical Determinism

In addition to the causal and theological forms of determinism, there is also logical determinism. Logical determinism builds off the law of excluded middle and holds that propositions about what agents will do in the future already have a truth value. For instance, the proposition "Allison will take the dog for a walk next Thursday" is already true or false. Assume that it is true. Since token propositions cannot change in truth value over time, it was true a million years ago that Allison would walk her dog next Thursday. But the truth of the relevant proposition is sufficient for her actually taking the dog for a walk (after all, if it is true that she will walk the dog, then she *will* walk the dog). But then it looks like no matter what happens, Allison will in fact take her dog for a walk next Thursday and that this has always been the case. However, it is hard to see how Allison's deciding to walk the dog can be a free decision since she must (given that the relevant token proposition *is* true and *was* true a million years ago) decide to walk him. In response to this problem, some philosophers have attempted to show that free will is compatible with the existence of true propositions about what we will do in the future, and others have denied that propositions about future free actions have a truth value, that is, that the law of excluded middle fails for some propositions. [For an introduction to these issues, see Finch and Warfield, (1999) and Kane, (2002).] If God is a being who knows the truth value of every proposition, this debate also connects with the debate over the relationship between [divine foreknowledge and free will](#).

From this brief survey, we see that free will touches on central issues in metaphysics, philosophy of human nature, action theory, ethics and the philosophy of religion. Furthermore, we've seen that there are competing views regarding virtually every aspect of free will (including whether there is, or even could be, such a thing). Perhaps this partially explains the perennial philosophical interest in the topic.

## 7. References and Further Reading

- Anglin, W. S. (1990). *Free Will and the Christian Faith* (Clarendon Press).
- Broad, C. D. (1952). "Determinism, Indeterminism, and Libertarianism," in *Ethics and the History of Philosophy* (Routledge and Kegan Paul).
- Chisholm, Roderick (1967). "He Could Have Done Otherwise," *Journal of Philosophy* 64: 409-417.
- Descartes, René (1998). *Discourse on Method and Meditations on First Philosophy*, 4th edition (Hackett Publishing Company).
- Ekstrom, Laura Waddell (1999). *Free Will: A Philosophical Study* (HarperCollins Publishers).
- Finch, Alicia and Ted Warfield (1994). "Fatalism: Logical and Theological," *Faith and Philosophy* 16.2: 233-238.
- Finch, Alicia and Ted Warfield (1998). "The Mind Argument and Libertarianism," *Mind* 107: 515-528.
- Fischer, John Martin (1984). "Power Over the Past," *Pacific Philosophical Quarterly* 65: 335-350.
- Fischer, John Martin (1994). *The Metaphysics of Free Will* (Blackwell).
- Fischer, John Martin and Mark Ravizza (1998). *Responsibility and Control: A Theory of Moral Responsibility* (Cambridge University Press).
- Frankfurt, Harry (1969). "Alternate Possibilities and Moral Responsibility," reprinted in Pereboom, (1997), pages 156-166.
- Frankfurt, Harry (1971). "Freedom of the Will and the Concept of a Person," reprinted in Pereboom (1997), pages 167-183.
- Ginet, Carl (1966). "Might We Have No Choice," in Keith Lehrer, ed., *Freedom and Determinism* (Random House), pages 205-224.
- Helm, Paul (1994). *The Providence of God* (InterVarsity Press).
- Honderich, Ted (2002). *How Free are You?*, 2nd edition (Oxford University Press).
- Kane, Robert (1998). *The Significance of Free Will* (Oxford University Press).
- Kane, Robert, ed. (2001). *Free Will* (Blackwell).
- Kane, Robert, ed. (2002). *The Oxford Handbook of Free Will* (Oxford University Press).
- Kane, Robert (2005). *A Contemporary Introduction to Free Will* (Oxford University Press).
- Lewis, David (1981). "Are We Free to Break the Laws?" *Theoria* 47: 113-121.
- McKay, Thomas and David Johnson (1996). "A Reconsideration of an Argument against Compatibilism," *Philosophical Topics* 24: 113-122.
- O'Connor, Timothy (2000). *Persons and Causes: The Metaphysics of Free Will* (Oxford University Press).
- Pereboom, Derk, ed. (1997). *Free Will* (Hackett).
- Pereboom, Derk (2001). *Living Without Free Will* (Cambridge University Press).
- Smilansky, Saul (2000). *Free Will and Illusion* (Clarendon Press).
- Strawson, Galen (1994). "The Impossibility of Moral Responsibility," *Philosophical Studies* 75: 5-24.

Strawson, Peter (1963). "Freedom and Resentment," reprinted in Pereboom (1997), pages 119-142.

Stump, Eleonore (2003). *Aquinas* (Routledge).

Van Inwagen, Peter (1983). *An Essay on Free Will* (Clarendon Press).

Widerker, David and Michael McKenna (2003). *Moral Responsibility and Alternative Possibilities: Essays on the Importance of Alternative Possibilities* (Ashgate).

## Author Information

Kevin Timpe

Email: [ktimpe@sandiego.edu](mailto:ktimpe@sandiego.edu)

University of San Diego

Last updated on March 31, 2006 | Categories: [Metaphysics](#)

Originally published on: March/31/2006

## Search the IEP

 submit

## Browse by Title

[A](#) [B](#) [C](#) [D](#) [E](#) [F](#) [G](#) [H](#) [I](#) [J](#) [K](#) [L](#) [M](#) [N](#) [O](#) [P](#) [Q](#) [R](#) [S](#) [T](#) [U](#) [V](#) [W](#) [X](#) [Y](#) [Z](#) [Browse by Topic](#)

▶ History of Philosophy

▼ Metaphysics & Epistemology

[Epistemology](#)  
[Metaphysics](#)  
[Mind & Coanitive Science](#)  
[Philosophy of Lanuaaqe](#)  
[Philosophy of Religion](#)

▶ Philosophical Traditions

▶ Science, Logic, & Mathematics

▶ Value Theory

## Recent Articles

[Reliabilism](#)  
[Santideva](#)  
[Time](#)

[Home](#)  
[About](#)  
[Staff](#)  
[Desired Articles](#)  
[Submit](#)  
[Volunteer](#)  
[Sitemap](#)

© [Copyright Information](#)