

Free Will and (In)determinism in “Hang the DJ”

Taylor W. Cyr

Forthcoming in Amber Bowen and John Anthony Dunne, eds., *Theology and Black Mirror*,
Fortress Academic; please cite published version.

In one of *Black Mirror*'s least pessimistic episodes, “Hang the DJ,” the main characters, Frank and Amy, are participating in a sort of futuristic dating service.¹ Participants in this service live on an idyllic compound and are assigned relationships of varying lengths by an algorithm called “the system.” The system’s ostensive reason for selecting matches is to learn from participants’ responses and ultimately determine the best possible match for every participant. Frank and Amy are initially set up by the algorithm with an expiry date of only twelve hours after they meet, and the two are left wishing they had more time to get to know each other.

Frank immediately gets matched with another partner, Nicola, and the expiry date for this match is a year. Frank and Nicola despise each other, and Frank clearly wishes he could have had more time with Amy. Meanwhile, Amy is assigned a series of matches with shorter expiries and becomes disillusioned with the process. Eventually Frank and Amy are matched again and are clearly very excited about this. Amy suggests that they not check the expiry date this time around, and Frank agrees. After some time, however, and partly because things are going so well and Frank is curious how long it will last, Frank checks the expiry date. His doing so without Amy checking at the same time causes the original time, several years, to shrink to a matter of hours. On their final day together, Frank is clearly aloof, and Amy eventually prods him until he confesses to having broken their agreement and shortened the length of the relationship.

¹ The series creator and writer of the episode, Charlie Brooker, calls it a “companion piece” to another episode, “San Junipero,” noting its “light and playful comic tone” in Charlie Brooker and Annabel Jones, with Jason Arnopp, *Inside Black Mirror* (New York: Crown Archetype, 2018): 272.

Despite this hiccup, Frank and Amy want to be together, and they do not connect with subsequent matches. The final day of their time in this dating service approaches, and Amy encourages Frank to rebel against the system and to flee the compound. When approached by a man with a taser who intends to stop them, Amy reaches out and touches the taser, and everyone besides Frank and Amy freezes, revealing that the two are part of a simulation—a possibility that had been floated in conversations in earlier scenes. Frank and Amy run to the edge of the compound where there is a wall with a very tall ladder, and as they are climbing the simulation around them begins to disappear. In the penultimate scene, the Frank and Amy we've been watching look around to find hundreds of other iterations of themselves, most of whom (nine hundred and ninety-eight out of one thousand) have also rebelled against the system and chosen each other. In the final scene, we learn that these simulated versions of Frank and Amy are part of a dating app. The ostensibly real Frank and Amy have used the app and discovered that they are a nearly perfect match. The episode ends with exchanged glances of hope for a relationship, even before they introduce themselves.

Like most episodes of *Black Mirror*, “Hang the DJ” raises a host of philosophical questions. In addition to the obvious ethical issues surrounding algorithm-matchmaking,² the very premise of the episode requires us, the audience, to assume controversial positions on the nature of consciousness, such as that there could be simulated consciousnesses. This invites us to ask about the connection between the simulated consciousnesses (“sims” for short) used by the software, on the one hand, and the users in the more fundamental level of reality (outside the

² Some such issues are explored in Aidan Powerm “Stop Me If You Think You’ve Heard This One Before: Relationships and Late Capitalism in ‘Hang the DJ,’” in T. McSweeney and S. Joy (eds.), *Through the Black Mirror* (Palgrave Macmillan, 2019) and Skye Cleary and Massimo Pigliucci, “Hand the DJ and Digital Dating: Should We Use Comptuers to Help Us Find Mates?” in D. K. Johnson (ed.), *Black Mirror and Philosophy: Dark Reflections* (John Wiley & Sons, Inc., 2020).

simulation): are these sims in some sense “versions” of the users, or are they merely qualitatively similar identities? And beyond these questions are questions about the limits of our knowledge about ourselves in our world given that, for all we know, we might ourselves be sims.

While there is much from this episode to explore, this chapter will explore something that has not yet been addressed in other work, namely the connection between “Hang the DJ” and questions about free will and determinism (or indeterminism, as the case may be). The topic of *free will* is important in both philosophy and theology, but it is worth mentioning that the expression *free will* is often used in different ways. For our purposes, it will be useful to distinguish *free will* from *autonomy*, the latter of which can refer specifically to having control over one’s own life without interference (or design) by another agent. It is this sense of freedom, of autonomy, that is suggested by the song played at the end of “Hang the DJ”—“Panic,” by The Smiths, from which the *Black Mirror* episode gets its title. When we hear the words of the song, which speak to the irrelevance of popular music for real life, the conclusion is that we should “Hang the DJ.” Certainly this episode asks us to consider the nature and value of autonomy, especially in the episode’s twist (when “Panic” is playing) and we discover the role that the software is playing in the main characters’ decision to date.

But in the traditional debates about free will, the term *free will* is used to refer to a kind of control over our conduct that we assume that we have. This control requires having the ability to do otherwise than what we actually do.³ For many free will theorists, the notion of free will is especially important because of its connection to moral responsibility. We ordinarily take

³ There are several great introductory books on free will, but for anyone interested in reading more about this debate I would recommend starting with Meghan Griffith, *Free Will: The Basics* (New York: Routledge, 2013).

ourselves and others to be morally responsible for much of what we do, and yet, plausibly, having free will is a necessary condition on moral responsibility. In addition, it is sometimes suggested that love requires free will, such that relationships not entered freely could not count as genuinely loving relationships. It is this sense of freedom (and control) that I will focus on here.

This chapter will proceed as follows: first, I will sketch some reasons for thinking that, if determinism is true, then no one has or exercises free will. One type of response to determinism's threat to free will is to accept the incompatibility of free will and determinism and to maintain that we nevertheless have free will. Theorists who endorse indeterministic accounts of free will are called *libertarians* in the free will debate (but please do not confuse them with political libertarians). Second, I will explain a bit more of the mechanics of libertarianism. Third, I will discuss an influential challenge to libertarianism that has come to be known as the "rollback argument." The mechanics of this challenge will resemble the plot twist of "Hang the DJ." Fourth, and finally, I will explore the episode's portrayal of the value of undetermined choice.

Determinism's Apparent Threat to Free Will

Determinism is sometimes defined as the thesis that there is, at any instant, only one physically possible future.⁴ Here's the basic idea: think of the laws of nature as a function, and think of a description of the state of the world at some time as an input into that function. (And by "state of the world" I mean everything about the world at that time, down to location and intrinsic properties of the tiniest microphysical particles.) The idea is that if the laws are deterministic, then from them as an output, and any input you would get, describes the state of the world at every subsequent time. In other words, from a description of the state of the world long before there were any human

⁴ See Peter van Inwagen, *An Essay on Free Will* (Oxford: Clarendon Press, 1983): 3.

beings, together with the deterministic laws, one could, in principle, deduce all of human history—even down to what you will choose to have for breakfast tomorrow. Technically speaking, if determinism is true, then propositions describing all of our actions are entailed by propositions expressing the laws of nature and propositions about the intrinsic state of the world long before we existed.

Let's pause to reflect on the simulated world of "Hang the DJ." Are the simulations programmed to work in a deterministic way, or are the sims' choices undetermined? We get an answer in the penultimate scene of the episode—the scene where the sims we have been watching appear in a space with other sims. We learn that the simulation of the relationship has been run a thousand times and that in all but two of those simulations the sims chose each other. A natural interpretation of this scene is that the sims' choosing each other is not determined but is nevertheless extremely likely, given their profiles. A different interpretation, consistent with the sims' being determined, is that the conditions the sims are placed into varies from iteration to iteration.

Now, some may regard the indeterminacy of "Hang the DJ" as a potential haven for free will, for it will seem obvious to some readers that determinism is an apparent threat to free will. One may even be tempted to construe the free will debate as the debate over free will versus determinism; obviously they are in conflict (so the thought goes), and thus we must choose one or the other but not both. It turns out, however, that to present the debate in this way is to presuppose a position on one of the main points of disagreement in the debate, namely whether or not free will is compatible with determinism. Compatibilists do not think we must choose between free will and determinism; their view is that these two things are compatible! I will now go on to explain

particular challenges for compatibilism—two arguments for *incompatibilism*—but one must keep in mind that not everyone is convinced by these arguments.

The first argument for incompatibilism I want to mention is the “consequence argument,” which may be summarized as follows.⁵ As we have seen, if determinism is true, then propositions describing all of our actions are entailed by propositions expressing the laws of nature and propositions about the intrinsic state of the world long before we existed. Now choose any action that you have ever performed or will perform, and call that action *A*.

1. If determinism is true, then you doing *A* is the consequence of the distant past and laws of nature.
2. For you to have the freedom to do otherwise than *A*, at least one of the following must be true:
 - a. You have the freedom to act in such a way that the past would have been different than it actually was.
 - b. You have the freedom to act in such a way that an actual law of nature would not have been a law.
3. You don’t have the freedom to act in such a way that the past would have been different than it actually was.
4. You don’t have the freedom to act in such a way that an actual law of nature would not have been a law.
5. If determinism is true, then you lack the freedom to do otherwise than *A*.

⁵ I am simplifying a conditional formulation of the argument along the lines of John Martin Fischer, *The Metaphysics of Free Will: An Essay on Control* (Oxford: Blackwell, 1994), but the *locus classicus* is Peter van Inwagen, *An Essay on Free Will* (Oxford: Clarendon Press, 1983): 55-105.

The first premise of this argument is meant to be an implication of the truth of determinism. If determinism is true, only one future is physically possible (i.e. consistent with the past and laws), and so what happens in that future, including our behavior, is the inevitable consequence of the past and laws. To do otherwise, then, would require the past or the laws (or both) to have been different, and this is what the second premise says. The third and fourth premises make claims about our abilities, specifically about our lacking the freedom to act in ways inconsistent with the actual past or actual laws, respectively. The third premise can be motivated by reflecting on how the state of the world in the distant past lies outside of our control. Similarly, the fourth premise can be motivated by reflecting on how what the laws of nature are lies outside of our control as well. If the foregoing is correct, then it looks like the truth of determinism would preclude our having the freedom to do otherwise than what we actually do.

In the introduction I mentioned the connection between free will and moral responsibility. It is common nowadays for philosophers working on free will to treat it as a necessary condition of moral responsibility. Since determinism threatens the freedom to do otherwise, then, we might ask whether that sort of freedom (the freedom to do otherwise) is the sort of freedom necessary for moral responsibility. And many theorists, such as those inspired by Harry Frankfurt, think that we can be morally responsible even if we lack the freedom to do otherwise.⁶ While a fascinating and ever-growing discussion, we can side-step this corner of the debate by looking at a different argument for incompatibilism—one that targets even those compatibilists who deny that moral responsibility requires the freedom to do otherwise.

⁶ Harry Frankfurt, “Alternate Possibilities and Moral Responsibility,” *Journal of Philosophy* 66 (1969): 829–839.

This second argument for incompatibilism is the “manipulation argument,” and the basic worry is that being causally determined by factors beyond one’s control (which compatibilists maintain is compatible with being morally responsible) looks relevantly similar to being manipulated, which many people take to undermine moral responsibility. The argument comes in various forms, but I will summarize a version developed by Alfred Mele (2006), which has come to be known as the “zygote argument.”⁷ Mele presents the following case involving a goddess, Diana, and an agent she creates, Ernie:

Diana creates a zygote *Z* in Mary. She combines *Z*’s atoms as she does because she wants a certain event *E* to occur thirty years later. From her knowledge of the state of the universe just prior to her creating *Z* and the laws of nature of her deterministic universe, she deduces that a zygote with precisely *Z*’s constitution located in Mary will develop into an ideally self-controlled agent who, in thirty years, will judge, on the basis of rational deliberation, that it is best to *A* and will *A* on the basis of that judgment, thereby bringing about *E*. (2006: 188)

With that case in mind, consider the following argument:

⁷ Alfred Mele, *Free Will and Luck* (New York: Oxford University Press, 2006). Another important version of the manipulation argument is Derk Pereboom’s “four-case argument.” See Derk Pereboom, *Living Without Free Will* (Cambridge: Cambridge University Press, 2001): 110-117 and Derk Pereboom, *Free Will, Agency, and Meaning in Life* (Oxford: Oxford University Press, 2014): 74-82.

1. Because of the way his zygote was produced in his deterministic universe, Ernie is not a free agent and is not morally responsible for anything.
2. Concerning free action and moral responsibility of the beings into whom the zygotes develop, there is no significant difference between the way Ernie's zygote comes to exist and the way any normal human zygote comes to exist in a deterministic universe.
3. So determinism precludes free action and moral responsibility. (2006: 189)

Given that it seems that any proposed compatibilist conditions on moral responsibility could be satisfied by Ernie after being created by Diana (he is self-controlled, responsive to reasons, doing what he most wants to do, etc.), the second premise looks quite plausible from the compatibilist's point of view (though, as one would expect, some do object to that premise). The first premise relies on our having a certain judgment about the case, namely that Ernie's moral responsibility is somehow undermined because of the way he was created, and it may be that the best way for compatibilists to respond to the argument is to deny the first premise.

Indeterministic (Libertarian) Free Will

Suppose you are convinced by one or both of the arguments summarized in the previous section (or suppose you have some other reason for thinking that free will and determinism are incompatible). If we were to discover that our world is deterministic, then we would know that we lack free will. But assuming that we don't know whether our world is deterministic, what would be required in order for us to have free will? What sort of agents would we need to be, and what sort of powers or abilities would we need to possess?

Some incompatibilists are not sanguine about our prospects. One subset of this group of free will skeptics—impossibilists—think that free will is impossible, and so of course not something we possess. Another subset—hard incompatibilists—wouldn't go so far as to make the impossibility claim but would say that we lack free will nonetheless. In any case, combined with the threat to free will from determinism, free will skeptics do not see *indeterminism* as a refuge for free will.

But the incompatibilists we are interested in here are those who think that we do have free will, i.e., *libertarians*, and they are united in taking free will and moral responsibility to require indeterminacy somewhere in the causal sequence leading to our behavior. Libertarians disagree about why determinism undermines free will, and there are two important strands of thought worth mentioning. Some emphasize determinism's threat to the freedom to do otherwise or to our having alternative possibilities, which they take to be essential to the freedom required for moral responsibility. Others emphasize determinism's threat to our being genuine sources of our behavior—the buck stopping with us, so to speak. In either case, for an action to be directly free and one for which an agent is directly morally responsible (direct in the sense that the agent's freedom and moral responsibility does not depend entirely on some earlier free action), that action must not be deterministically caused. Instead, it must be the case that, holding fixed everything about the past right up to the time of action, and holding fixed the laws of nature, another course of action (or at least some alternative—perhaps not acting at all) was possible for the agent.

Let's consider a concrete case, one taken from "Hang the DJ." Suppose that Frank's decision to look at the expiry date of his match with Amy is an undetermined decision. That means that, holding fixed the laws of their universe, and holding fixed everything that has taken place in it right up to the moment of decision, an alternative to checking the expiry date is possible. Whether

one emphasizes the need for alternative possibilities or the need to be the genuine source of one's actions, it looks like Frank's decision being indeterministically brought about helps to assuage the worries about determinism's threat to free will.

Thinking back to the arguments for incompatibilism summarized in the previous section, we can see why libertarians would think that indeterministic causation may leave room for free will. First, if one takes the freedom to do otherwise to be crucial for genuine free will and moral responsibility, then indeterminacy at the time of action would seem essential for freedom and responsibility. If my performing a certain action is not determined by the past and laws, then my action is not just an inevitable consequence of factors beyond my control, and this may seem relevant to the control I exercise in acting. Second, if one takes determinism's main threat to be a threat to our being genuine sources of our behavior, especially given the similarities between ordinary determinism and manipulation, then (again) indeterminacy at the time of action would seem essential for freedom and responsibility.

The Rollback Argument

Perhaps the most important (and certainly the most widely discussed) challenge to libertarianism about free will is at its heart a worry about luck, randomness, or chance. Here too there are different ways of spelling out the worry, but the basic idea is that, given the libertarian's requirement that directly free actions be undetermined, it seems that meeting that requirement makes it a matter of luck (or outside the agent's control) that the agent acts in one way rather than another. The agent's acting in that way is not settled beforehand by anything in the agent, and for every action satisfying this condition there is some alternative scenario where everything was the same right up to that same point but in which the agent did something else instead. That makes undetermined actions

look more like a chancy outcome than a real expression of agency. (It is worth noting that, since most compatibilists take free will to be compatible with indeterminism as well as determinism, including indeterminism at the very time of action, this problem of luck presents a challenge for those compatibilists views too. The reason that the worry is typically associated with libertarianism in particular, though, is that the problem arises from what the libertarian—and not the compatibilist—takes to be a necessary condition of free will.)

One articulation of this worry that has been widely discussed is Peter van Inwagen's (2000) "rollback argument."⁸ Consider the following case:

Let us suppose undetermined free acts occur. Suppose, for example, that in some difficult situation Alice was faced with a choice between lying and telling the truth and that she freely chose to tell the truth—or, what is the same thing, she seriously considered telling the truth, seriously considering lying, told the truth, and was able to tell the lie she had been contemplating. And let us assume that free will is incompatible with determinism, and that Alice's telling the truth, being a free act, was therefore undetermined. Now suppose that immediately after Alice told the truth, God caused the universe to revert to precisely its state one minute before Alice told the truth (let us call the first moment the universe was in this state 't1' and the second moment the universe was in this state 't2'), and then let things "go forward again." What would have happened the second time? What would have happened after t2? Would she have lied or would she have told the truth? Since Alice's "original" decision, her decision to tell the truth, was undetermined—since it was

⁸ Peter van Inwagen, "Free Will Remains a Mystery," *Philosophical Perspectives* 14 (2000): 1-19.

undetermined whether she would lie or tell the truth—, her “second” decision would also be undetermined, and this question can therefore have no answer... (van Inwagen 2000: 14)

Alice’s action of telling the truth meets the libertarian’s requirement of being undetermined, but in virtue of meeting that requirement the action appears a matter of chance. Van Inwagen continues:

Now let us suppose that God *a thousand times* caused the universe to revert to exactly the state it was in at t_1 (and let us suppose that we are somehow suitably placed, metaphysically speaking, to observe the whole sequence of “replays”). What would have happened? What should we expect to observe? Well, again, we can’t say what would have happened, but we can say what would probably have happened: sometimes Alice would have lied and sometimes she would have told the truth. As the number of “replays” increases, we observers shall—almost certainly—observe the ratio of the outcome “truth” to the outcome “lie” settling down to, converging on, some value. We may, for example, observe that, after a fairly large number of replays, Alice lies in thirty percent of the replays and tells the truth in seventy percent of them—and that the figures ‘thirty percent’ and ‘seventy percent’ become more and more accurate as the number of replays increases. But let us imagine the simplest case: we observe that Alice tells the truth in about half the replays and lies in about half the replays. If, after one hundred replays, Alice has told the truth fifty-three times and has lied forty-eight times, we’d begin strongly to suspect that the figures after a thousand replays would look something like this: Alice has told the truth four hundred and ninety-

three times and has lied five hundred and eight times. Let us suppose that these are indeed the figures after a thousand replays. Is it not true that as we watch the number of replays increase, we shall become convinced that what will happen in the *next* replay is a matter of chance? (van Inwagen 2000: 14-15)

Even if we suppose that Alice has reasons for telling the truth and reasons for lying, and that either course of action will be explicable in terms of her agency (and her reasons for action), the rollback scenario highlights that it is odd to think of her behavior as free. No one exercises free will in rolling a six (with a fair die) even if they freely roll the die. Even for coin-flips, one may freely flip the coin, but no one brings it about that it lands “heads” through an act of free will. It seems that the luck involved in these cases is inimical to control, or freedom. And yet the rollback scenario highlights that, for agents whose behavior is undetermined, which course of action they take seems to be a matter of luck.

Now, while I disagree with van Inwagen and think that the libertarian can solve this problem without appealing to mystery, we need not explore potential avenues of response here. What is of interest for our purposes, and what should be quite obvious at this point, is that the mechanics of the rollback argument closely resemble the “Hang the DJ”’s *denouement*. Even if we interpret the various sims portrayed in the episode’s penultimate scene as different instantiations of the real Frank and Amy’s psychological profiles (rather than, say, as the very same sims being rolled back hundreds of times), it is suggested that these sims are put in the same circumstances and that, in all but two of the one thousand iterations, the sims of Frank and Amy choose each other. What this set of replays indicates, then, is that Frank and Amy are not determined to choose each other but are nevertheless very likely (perhaps nearly determined) to

do so. And the portrayal of these replays poses certain questions for us concerning the value of undetermined choice, which we will consider in the next and concluding section.

The Value of Undetermined Choice in “Hang the DJ”

I mentioned at the outset that one reason for caring about having free will had to do with the idea that loving relationships require free will. Perhaps one cannot be in a genuine relationship of love if one did not enter into that relationship freely. For the libertarian in particular, motivated by the worries codified by the consequence and manipulation arguments, it may seem that a relationship’s being determined would very obviously preclude its being a relationship of love. If determined, the relationship will appear the mere consequence of the distant past and the laws of nature, and those in the relationship may even seem not too dissimilar from puppets. If entering into the relationship is undetermined, however, then genuine love may seem a real possibility, for the people in the relationship were not bound, in any sense, to enter into it.

What is interesting and, I think, very suggestive about “Hang the DJ,” though, is that the closer to being determined to choose each other Frank and Amy are, the more positive their relationship is portrayed. The fact that their sims have chosen each other rollback after rollback (or in hundreds of simultaneous simulations) strikes us as *good*, at least insofar as the knowledge of the likelihood of their choice of each other is helpful in the context of finding a match. Now, one might take this suggestion in either of two directions, depending on whether one took free will to be compatible or incompatible with determinism. The compatibilist can simply accept that determined (or nearly determined) agents nevertheless possess free will, and so they can readily maintain that loving relationships require free will. For the libertarian (who is an incompatibilist),

however, the portrayal of Frank and Amy suggests that it is not very important that we enter into relationships freely—perhaps this sort of undetermined choice is not particularly valuable.

A related issue concerns not our relationships with one another (other human beings) but rather with God. Many theists are attracted to libertarianism about free will partly due to worries about the (im)possibility of our genuinely coming to love God if our coming to do so is determined. If we come to see that loving relationships need not be entered freely, however—or, if we are compatibilists, that they can be entered freely even if determined—some of the motivation for denying that God could determine us to respond in love will evaporate. Of course, this does not mean that it does not matter the means by which we come to be in a loving relationship with God. In Frank and Amy’s case, they choose one another not as a result of external force or coercion but rather because they really like each other. Similarly, so long as a human being is responding to God in love because of features of God—perhaps goodness, beauty, grace, and the list goes on—there is no reason to think that being determined to respond in love is anything like the result of force or coercion.