

MIT OpenCourseWare
<http://ocw.mit.edu>

24.120 Moral Psychology
Spring 2009

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.

XIII Free Will I

Free will presents a problem for metaphysics, philosophy of mind and ethics. Indeed, it presents such a serious problem exactly because its repercussions have such ethical importance, and stretch right to our very conception of ourselves.

The Problems

Firstly: there are set of specific reasons for thinking that we are neither so free in forming our desires, nor in acting on them, as we might think. Consider the correlation between people's tastes and choices, and their social class and upbringing. (One of the best predictors of how an individual will vote is how their parents voted.) It doesn't matter whether this is nature or nurture, the upshot is still the same. Consider the effectiveness of advertizing. More particularly, there is a wealth of psychological literature showing that we are less free in our choices that we might think. Women choosing a pair of tights will consistently choose the one on the right of the display. (Similar findings apply to men!) But these considerations seem to undermine both our freedom to form our desires, and our freedom to act upon them.

Secondly, and more generally: we have a picture of the world that is given to us by science according to which we are part of the natural world. But the natural world is governed by deterministic causal laws. So everything we do, along with everything else, is deterministically caused by what went before. This is the thesis of determinism.

- PI If determinism is true, then every human action is causally necessitated
- P2 If every action is causally necessitated, no one could have acted otherwise
- P3 One only has free will if one could have acted otherwise
- P4 Determinism is true
- C No one has free will

This argument is clearly valid. So disagreements will focus on whether or not it is sound; and if it isn't on which premise(s) should be rejected. Recall the standard terminology. Hard determinists accept the soundness of the argument and so embrace its conclusion. Libertarians deny its conclusion, and do so by denying P4. (Note that it is not enough just to deny determinism. We have to say what to put in its place. And it is quite unclear what could play the role. Certainly thinking that events happen randomly, as quantum mechanics is sometimes held to entail, will not do the job.) Compatibilists deny the conclusion and accept P4—they want to hold that determinism and free will are *compatible*—and so standardly want to reject one of the other premises; typically P2 or P3 (or both). But some positions that look like compatibilism turn out, on closer examination, to be arguing for the compatibility of determinism with our normal practices of holding people responsible. So such positions seem to be able to accept the soundness of this argument. They are compatibilists not about freedom and determinism, but about responsibility and determinism.

To get a clearer view of quite what is at issue here we need to step back a little. Why should the conclusion seem so threatening? There are two distinct sorts of consideration that make it so. One concerns the phenomenology of freedom, the other concerns our ascriptions of responsibility.

The Phenomenology

It seems as though we are free to choose. Indeed, to say that it seems that way is to understate the point badly. There are few things more sure than that we are free to choose. As Dr. Johnson famously said: 'Sir, we know our will is free, and there's an end on't'. Imagine a trivial example: you are faced with a choice of dessert. You are very partial to both; but you can only have one or the other. You feel your self with the choice between them; you feel as though it is up to you to decide. You lean first towards one, then the other; you might wonder which one you will chose. Finally, you make a choice. It feels as though you could have chosen differently. And, until you act on your decision, it feels as though you could still change your mind. Now imagine some more important cases: you are deciding which university to apply to, which courses to take, which person to spend the rest of your life with.

These considerations are based on a certain phenomenology: we have an experience that things feel a certain way. Experiences aren't themselves propositions, so they can't directly be consistent or inconsistent with the conclusion of an argument. But we do think that experiences reveal the world to be a certain way. In just the way that our visual experience tells us that, for instance, we are looking at a red cube, so our experience of acting tells us that we have free will. But what exactly does it tell us? There are various possibilities (in increasing order of strength):

- (i) that we really do make choices, and that these are determiners of how we act;
- (ii) that we could have acted otherwise;
- (iii) that our choices are not caused by anything else.

If this is right, then the conclusion of the argument is worrying. Quite how worrying it is depends on which of these possibilities we take to be revealed by our experience.

Responsibility

What sorts of things do we ascribe responsibility to? They need to be an agent with a will that they can exercise. Unless we are badly anthropomorphising, we don't blame a car that has broken down. Nor do we blame someone for not helping when we find that they were locked in a room, unable to come to our aid. Moreover they need to have the right kind of control over their will; we don't blame a dog, or a small child for eating some cakes that we have left out in the same way that we would blame a roommate for doing the same thing. And more than that, they need to have got the desire in the right way; we don't blame someone for doing that which they were induced to do under hypnosis; we don't blame someone for doing what they are forced to do as a result of a compulsion (e.g. the kleptomaniac; note though that here we seem to be getting closer to the first worry). It seems then that we are to blame for that which we freely do as a result of desires that we autonomously form. So there seem to be at least two aspects of free will that are essential to our practice of ascribing blame and responsibility: freedom to form our desires in the right sorts of ways, and freedom to act upon them.

But, again, if this conclusion is right, then the consequences for our ordinary moral practice seem to be devastating. Since, as we have seen, we only blame people, feel resentment towards them, etc., if we think that their actions are freely performed, then if no action is freely performed, we should give up on these attitudes. This is indeed the conclusion that some people have drawn, for instance psychologists like B. F. Skinner. But that seems to many an abhorrent option.

Characterizing Free will

What is it for someone to have free will? The natural way is to think that, although they act a certain way, they could have done otherwise. And a natural way to understand that is in terms of possibility: someone could have done otherwise if it is possible that they did otherwise. But if we understand possibility in the standard possible worlds sense, that is much too weak. It is of course possible that you did otherwise, in the sense that there is a possible situation perhaps with a very different history to the actual one, in which you did otherwise. No determinist needs to deny that—determinists aren't committed to saying that there is only one way that the world could have been. They are not fatalists. The worry is rather that, given the laws of nature, and the facts about the world just before you acted, there is no possibility that you did otherwise.

A first argument for Compatibilism: Hobbes' Solution

A FREE-MAN, is he, that, in those things which by his strength and wit he is able to doe, is not hindred to doe that he has a will to ... from the use of the word *Free-will*, no Liberty can be inferred of the will, desire or inclination, but the Liberty of the man; which consisteth in this, that he finds no stop in doing what he has the will, desire or inclination to doe. (*Leviathan*, Ch XXI)

How could we use these considerations to refute the argument that is given above, viz:

- P1 If determinism is true, then every human action is causally necessitated
- P2 If every action is causally necessitated, no one could have acted otherwise
- P3 One only has free will if one could have acted otherwise
- P4 Determinism is true
- C No one has free will

One response would be to simply deny P3; free actions are those that stem from our desires, there is no need for it to be true that we could have done otherwise. But don't we want to preserve the idea that we could have done otherwise? How might this be done? A first try might be: to say that I *could* have acted otherwise is to say that I *would* have acted otherwise if I had *desired to do otherwise*. (G. E. Moore gives a closely related analysis: I would have acted otherwise if I had *chosen to do otherwise*.)

Compare windvanes: a windvane is free in so far as it points the way the wind is blowing. It isn't free to point to the East whilst the wind is blowing from the North; its freedom consists in the fact that if the wind had been blowing from the East, it would have pointed to the East.

Now P3 is properly understood as :

- P3* One only has free will if one could have acted otherwise (i.e. if one would have acted otherwise if one had desired to act otherwise).

But then P2 is straightforwardly false; our actions could be causally necessitated whilst it is true that we could have done otherwise, since we would have done otherwise if we had desired otherwise, in which case our actions would have been necessitated to have been different.

(Similarly, the movements of the windvane are causally necessitated; but it could have pointed in a different direction, since the wind could have been blowing from a different direction, in which case it would have been causally necessitated to have pointed in that direction.)

Problems

First, what about things that I know to be impossible? Isn't it true that I act freely when I decide not to fly by flapping my arms? As it happens I don't want to fly by flapping my arms. But suppose though that I did want to fly by flapping my arms. Would I do so? No. I know it wouldn't work. But then the analysis suggests that my decision not to flap my arms is not a free one, because it's not true that I would have acted otherwise if I had desired to do otherwise. (This is a general problem with any account that understands being free in terms of *being able to get what we want*. We don't think that the fact that certain things are impossible to get is a restriction on our free will.) Hobbes avoids this problem by limiting the account to those things that the agent 'by his strength and wit is able to do'; but this risks making the account circular. For what is the talk of ability here if it isn't talk of what one is free to do? Perhaps Hobbes can, and should, avoid the circularity by making the crucial clause concern what we *believe* we are able to do.

Second: doesn't the account put the weight back onto my desires. For what if my desires were not free? Suppose they were implanted in me by hypnosis or whatever. Then we wouldn't think that my act was free even if it were true that had I desired otherwise I would have acted otherwise. We might try reworking the argument by substituting 'choose' or *somesuch* for 'desire'. But parallel problems will arise: what if I couldn't have chosen otherwise? And now the worry is that if we try to give another conditional account of what it is to desire, or choose, or whatever, freely, we will be in a regress. Hobbes just rejects this worry; he says that free will concerns only our actions, not our wills, desires or inclinations. But many have thought that this is just to miss the point. This is the force of the consequence argument that we shall look at next session.

Third: is it even true that 'I could have done otherwise' means the same as 'I would have done otherwise if I had desired to do so'? Is it true that 'I could have shot him' means the same as 'I would have shot him if I'd wanted to'? In general, statements of possibility are not normally best understood as counterfactuals.

Frankfurt's Account of Freedom

First: Distinguish amongst your desires, between those that are *effective*, and those that are not. An effective desire is one that actually issues in action. Thus you might both want to go to Mexico, and go to France. If you go to France, that is the desire that is effective. Note that you can both desire to do something and desire to not do it. But only one of these can be effective. Frankfurt identifies a person's *will* with their effective desires.

Second: Distinguish, amongst your desires, between first order desires and second order desires. A first order desire is a desire for anything other than a desire; a second order desire is a desire for a desire. So, for instance, you might have a first order desire to smoke a cigarette, and a second order desire that you desire not to smoke a cigarette. A second order desire might or might not be a desire that its corresponding first order desire be effective. Thus I might wish that I wanted to give all my money to charity, since I might think that having such a desire

would show me to be an excellent person; but I might nonetheless not actually want that desire to be effective. I might think it would be terrible if I actually did give all my money away. But when a person does want the first order desire to be effective, when they want it to be their will, Frankfurt calls this a *second order volition*.

Some creatures have no second order volitions. Frankfurt calls such creatures *wantons*. He claims that non-human animals and children are all wantons in this sense. He contrasts wantons with persons, who do have second-order volitions. (This seems a bit stipulative; wouldn't we say that an adult human being who had no volitions was still a person?) As an example, compare two sorts of drug addicts, a wanton and an unwilling person. The person wants not to want the drug; but his craving is too strong. In contrast the wanton has no desires about his desire; he is indifferent to being an addict. (Note that the wanton can still have contradictory desires; he might both want and not want the drug; the important thing is that he has no preference as to which of these desires triumphs.)

Now Frankfurt is in a position to define freedom. He thinks there are two aspects. Firstly, there is the aspect we have already looked at: a person's actions are free in so far as they stem from their desires; that is, if they had desired differently, they would have acted differently. But for Frankfurt this is merely freedom of *action*. In addition a person has freedom of *will* (i.e. their effective desires are free) in so far as they can control their desires. That is, a person has free will just in case they have second order volitions, and they can bring their first order desires into line with them. (Frankfurt thinks that a person can lose their freedom of action and still have freedom of will; for instance if they do not realize that they cannot act they might still go on freely controlling their will in accord with their second order volitions.) So we can see why the cases we looked at earlier aren't cases of free will. The dog and the small child don't have free will, because they are wantons; they lack second-order volitions. The unwilling drug addict lacks free will for a different reason; he has a second order volition, but it does not control his effective desire. The kleptomaniac is similar. Similarly you do not have free will if you act on a desire that is implanted in you by hypnotism when you have a second order volition not to have that desire. (Although presumably if you do want to have that desire—perhaps you paid a hypnotist to get you not to want cigarettes because you thought that was the best way of quitting—then the act is still one performed from free will.)

Worries About Frankfurt's First Account

In pushing the problem up one level (from first order desire to second order) has Frankfurt really solved it? Two worries: (i) are second order desires really so special? Consider Jeanette Kennett's example of wanting to want strawberries (because it would be socially useful). (ii) Suppose that I implanted a second order desire in you by hypnosis. Then surely you wouldn't have free will if you got your desires to conform to that; but Frankfurt's account seems to have the consequence that you would. Of course, we could always then insist that there must be a third order desire; but suppose that too were implanted by means of the microchip, or by hypnotism? Whatever level we go to, we can wonder whether that desire is held freely; and then we seem to need to go up another level. We are in a regress. Frankfurt's response to worries like this is to say that there are certain desires that we *identify with*; and that free will consists in getting our desires into line with these. But what makes something a desire that we identify with? We have seen that it isn't just a matter of *level*. So it looks as though Frankfurt has left the central issue unanswered.