

Freewill III: "Deep Self" Compatibilism

I. Recap: Compatibilism or "Soft" Determinism, first attempts

STANDARD HARD DETERMINIST ARGUMENT

1. Whatever happens is determined by prior events. (Determinism)
2. I act freely iff I am able to act otherwise. (Avoidability Analysis of Freedom)
3. If my action is determined, I am unable to act otherwise.
4. So, I never act freely.

Hypothetical Choice Compatibilism:

I act freely iff I would have acted otherwise had I chosen to.

Internal sources of action ("second try") Compatibilism (See Stace: "Acts freely done are those whose immediate causes are psychological states of the agent.")

I act freely iff

- (i) the way I act is the result of my choice,
- (ii) the immediate causes of my choice are "internal" to me.

"No Coercion" + "No-Duress" Compatibilism

I act freely iff

- (i) the way I act is the result of my choice,
- (ii) the immediate causes of my choice are "internal" to me,
- (iii) my choice is not the result of coercion or duress, and
- (iv) I would have acted otherwise had I chosen to.

But one problem is that it isn't clear why certain constraints on my choices are compatible with freedom and some not – why are ordinary cases of coercion and duress a problem and other equally powerful determinants of my action, e.g., my character and my upbringing, not? Initial reflection on our practices suggests that we are keen to capture the following in our analysis:

Free

Ordinary deliberate action

Unfree

Physical constraint or coercion

Behavior caused by psychosis or phobias

Behavior caused by addiction

Behavior caused by brainwashing, hypnosis, indoctrination (?).

The key question before us is: What is distinctive of the kind of coercion and duress that disrupts our freedom? Condition (iii) doesn't help us answer this question, but assumes we have an answer. In effect, condition (iii) just states

iii*) my choice was freely made.

But we have no reason yet to think that this condition is *compatible* with determinism. In fact, if we have to rely on (iii*) then we've really made no progress in our understanding of freewill.

II. Frankfurt's "Deep Self" Compatibilism

Thus far we've worked with the background assumption (we've called it the "avoidability condition on freedom") that

I act freely only if I am able to act otherwise.

Is this an assumption we should question? Suppose you are in a room with everything you've always wanted (your favorite entertainment, friends, food, etc.); if anyone tried to remove you, you'd fight to stay. Suppose then someone locks the doors and windows. You can't leave, but you also don't want to. Are you unfree? Whether you could have left the room if you wanted seems irrelevant to the issue of whether you stay of your own free will. This suggests that the real issue is not whether one is constrained, but whether one's actions are based on choices that are in some important sense "one's own". In other words, this sort of example suggests that the "avoidability condition on freedom" isn't as important as we initially thought.

If we're keen to make sure our desires are "our own" and aren't compelled--just as we were concerned our actions were "our own" and not compelled, maybe we should push the analysis back a step:

I act freely iff my action is caused by a free desire or choice.

So similarly:

I choose/desire freely iff my choice/desire is itself caused by further choices/desires.

Does this help? Can we make sense of choosing our choices, or desiring our desires?

It may sound odd, but there is an insight here. Not all of our desires hang together as nicely as we might like. In fact most of the time we have within us any number of conflicting desires: the desire to finish our work and the desire to go to bed; the desire to tell your friend the painful truth and the desire to lie to him; and so on. For a particularly vivid example consider the addict who feels a powerful urge to take the drug but also, as seems natural, a strong desire to refuse it and go clean. The question is, which of these desires is she going to act on? Which desire will be effective? Which of these desires is going to express itself in actual behavior?

So far we've been talking about desires to act in certain ways, for short ordinary desires. The question was, when two such desires conflict, what determines which of them wins out and controls the person's actions? But in addition to our ordinary desires we have reflective desires: desires to have certain ordinary desires and not others, or at least to have certain of those desires be the ones that actually control our behavior. According to Frankfurt, if one has a desire for one's first-order desire to be effective, i.e., for it to be one's will, then this is a *second order volition*. Some creatures have second-order volitions, and some don't (children? Non-human animals? Thus the addict may well have the reflective desire to lose her desire to take the drug, or, failing that, to have her desire to refuse the drug be the one that controls her behavior.

So far so good, but what does this have to do with the question of freedom? Compare two heroin addicts. Each has a desire to shoot up and also a desire not to shoot up. Those are their ordinary (or first-order) desires. Further each has a reflective desire to act on the desire to abstain rather than the desire to take the drug (this is a second-order desire).

First-order desires: Desire to take the drug & desire to abstain

Second-order desire: Desire to act on *the desire to abstain*

However, suppose only one succeeds in carrying out this reflective desire, hence in abstaining the drug; let's call the one who abstains: *Iquit*. The other one would like to be acting on his desire to abstain but finds that he is incapable of this; the desire to take the drug remains the operative one. This second addict, let's call him *Noquit* seems like a perfect picture of an unfree will: his operative desires are not the ones he would like to be operative. This suggests that the person with true freedom is the one who enjoys a certain kind of harmony between the first-order desires he wants to be operative and the ones that are actually operative. In other words, freedom is a matter of having second order volitions and being able to get one's desires and actions to conform to them.

Susan Wolf, in an essay on these issues, argues that Frankfurt's view is one variant of a broader strategy. (See *RR*, pp.457-467) The guiding idea is that in addition to the passing desires and choices we make day to day, there is a deeper self in each of us who embodies our values, our primary concerns, our fundamental identity; and what's important for freedom is that this self is the one who is in the driver's seat, i.e., this self is determining what we do. Wolf says:

All [of the authors in question] agree that if we are responsible agents, it is not just because our actions are within the control of our wills, but because, in addition, our wills are not just psychological states in us, but expressions of characters that come from us, or that at any rate are acknowledged and affirmed by us. For Frankfurt, this means that our wills must be ruled by our second-order desires; for Watson, that our wills must be governable by our system of values; for Taylor, that our wills must issue from selves that are subject to self-assessment and redefinition in terms of a vocabulary of worth....[in other words]the key to responsibility lies in the fact that responsible agents are those for whom it is...the case that their wills are within the control of their selves in some deeper sense. (*RR*, 459)

SO:

Deep-Self Compatibilism (Harry Frankfurt and others)

Basic idea: Freedom is a matter of having an integrated self that governs how one acts.

I act freely iff my actions are in the control of choices endorsed by my "deeper self" i.e., iff my actions are controlled by the desires that I want to be operative, the ones I identify with.

Compatibility: As in the previous case: as long as the causes of the action and the choices are of the right sort (in this version, they come from the deep self) the action is free; this does not require that the choices are uncaused. In this case, moreover, there is no suggestion here that avoidability is required for freedom.

Problem cases: weakness of will, early indoctrination, brainwashing, young children?

- Do we want to say that in cases of weakness of will we are unfree? Shouldn't I sometimes be held responsible for actions done from such weakness?
- Plausibly, children don't have the capacities for reflective desires that adults have, and in some sense lack a "deep self". Does this mean they aren't capable of free action?
- Insofar as one's "deep self" is still determined, does this account allow us to be genuinely free? Who, after all, or worse, *what* is responsible for this deeper self? The problem just seems to be pushed back--if we aren't responsible for the deeper self, then how can we be really free? But

these demands threaten to push us back ad infinitum--or it would seem we must create ourselves from scratch.

- Does it matter whether one's "deep self" itself issues from "wholesome" or "unwholesome" external causes, e.g., problematic upbringing, education, etc.

Consider a case from Susan Wolf: JoJo is the son of Jo, an evil and sadistic dictator. JoJo is brought up to have fundamentally flawed values--that he endorses and identifies with. What's wrong with this picture? Why does JoJo seem unfree? Consider the legal definition of insanity:

A person is sane if (1) he knows what he is doing, and (2) he knows that what he is doing is, as the case may be, right or wrong.

As Wolf glosses it,

...the desire to be sane is a desire to be controlled (to have, in this case, one's beliefs controlled) by perceptions and sound reasoning that produce an accurate representation of the world, rather than by blind or distorted forms of response. The same goes for the second constituent of sanity--only, in this case, one's hope is that one's values be controlled by processes that afford an accurate conception of the world. Putting these two conditions together, we may understand sanity, then, as the minimally sufficient ability cognitively and normatively to recognize and appreciate the world for what it is. (RR 10th edition, 441)

"Sane Deep-Self" Compatibilism (Susan Wolf)

I act freely iff my actions are controlled by choices endorsed by my "deeper self," & that deeper self is sane: I am able to appreciate what I am doing, and whether it is right or wrong.

Basic idea: Freedom is a matter of having an integrated self governing how one acts, when that self has the capacity to make revisions/improvements in its beliefs, desires, values, etc. in light of information about the world (specifically information about what's right and wrong); in short, the integrated self must also have the capacity for self-correction.

Compatibility: Same as Frankfurt case.

Problem cases: actions out of sync with our "deep self", moral monsters

Note that on Wolf's view there are two ways to be "unfree"--to be insane, or to have your sane deep self not controlling what you do/will. But if these are cases where we aren't free, they are also cases where we aren't responsible. So the first issue is to consider whether we want to say that those who are "insane" are not responsible; and whether we want to say that those whose deep selves aren't in control aren't responsible...

- Do we really want to let moral monsters off the hook by denying them freedom?
- In cases of weakness of will, are we unfree?
- Can we be confident that *our* system of right and wrong is objective, and so makes us sane, and so makes us free? Might we be misguided about right and wrong and so unfree?
- Does this really get us what we want? Let's suppose I'm sane and so have the ability to see what is right and to do it, and to see what is wrong and avoid it. But in each case my abilities and their exercise in particular cases is completely determined--since after all, Wolf's view and

all compatibilist views claim that freewill is compatible with determinism. But if that's the case, am I really free? Should I get credit for what I've done (or not done)?

Other resources:

A webpage on freewill by Richard Holton (a Professor of Philosophy at MIT, who used to be at the Univ. of Edinburgh):

<http://homepages.ed.ac.uk/rholton/freewill/fwhome.html>