

# A Compatibilist Account of Free Will and Moral Responsibility

If Frankfurt is right, he has shown that moral responsibility is compatible with the denial of PAP, but he hasn't yet given us a detailed account of free will. When and how are we free, even if determinism is true? In his follow-up article, he attempts to supply answers.

## 1. Defining Terms: First, some terms:

- **First-Order Desire:** This is simply a desire to do something. Frankfurt expresses this as "A wants to X" (where X is an action).

Note that first-order desires can be "effective" (meaning that they move you to act) or "non-effective" (meaning that they DON'T move you to act). For instance, I had an effective desire to do a lecture on Frankfurt. I also have a desire to eat an entire bag of Reese's cups, but this desire has (so far) been non-effective. They can also compete. I have a desire both to take a trip to Hawaii and Alaska. If I end up going to Hawaii, it is that desire that ends up being effective.

- **Will:** When I have an effective first-order desire, this is my "will". I will myself to do something, and I do it. He writes,

"To identify an agent's will is either to identify the desire (or desires) by which he is motivated in some action he performs, or to identify the desire (or desires) by which he will or would be motivated when or if he acts. An agent's will, then, is identical with one or more of his first-order desires. But the notion of the will, as I am employing it, is not coextensive with the notion of first-order desires. ... Rather, it is the notion of an *effective* desire—one that moves (or will or would move) a person all the way to action." (8)

- **Second-Order Desire:** This is a desire ABOUT one's (first-order) desires. We can express this as "A wants to want to X" (where X is an action).

Note that my second-order desires can be either for effective or non-effective first-order desires. For instance, I want to want to exercise more—and I want to want this in a way that will be effective (i.e., where I will actually exercise).

But, Frankfurt points out that it doesn't HAVE to be this way. He describes a situation where a psychologist is studying the effects of addiction. He wishes he could "get inside the head" of the addict. In that sense he *wants to want to take drugs*. But, he doesn't actually want this in a way that it would be effective. He wants to have the desire for drugs so that he can better understand the addict—but he doesn't actually want to follow through and take drugs.

- **Second-Order Volition:** When I have a second-order desire to have a first-order desire that is EFFECTIVE, this is my (second-order) "volition".

THESE are the things that separate us as importantly different from other animals:

"Now it is having second-order volitions, and not having second-order desires generally, that I regard as essential to being a person." (10)

One more technical term:

- **Persons:** Persons are beings who have second-order volitions.

Note also that being a person requires being RATIONAL. In order to have second-order volitions, you need to be capable of *rationally reflecting* upon yourself and your desires. We are peculiar in that we can sort of "step back" and take a "bird's eye view" of our desires, and wish for some of them to be effective. You don't just WANT to study hard, you look "down" at that desire and APPROVE of it. You WANT to WANT to study hard—and in a way that will actually result in ACTION (i.e., actual studying). Animals don't seem to have this capacity. Frankfurt calls them "**wantons**". These are any creatures that lack second-order volitions.

"The essential characteristic of a wanton is that he does not care about his will. His desires move him to do certain things, without its being true of him either that he wants to be moved by those desires or that he prefers to be moved by other desires." (11)

**2. Frankfurt's Proposal:** Only "persons" have free will, Frankfurt says. And therefore, (since moral responsibility requires free will) they are the ONLY sorts of creatures that are capable of being morally responsible for their actions. Only "persons" are what we'd call "moral agents", deserving of praise and blame.

Putting all of this together, Frankfurt makes the following claims:

1. A is morally responsible for doing X if and only if A does X freely.
2. A does X of her own free will (i.e., she freely does X) if and only if doing X meets the following criteria:
  - (i) A has a desire to X
  - (ii) A reflects on whether the *desire to X* is a desire by which she wants to be motivated
  - (iii) A answers "yes"; i.e., she reflectively endorses/identifies with the desire to X
  - (iv) The desire to X moves A to perform action, X

Thus,

“It is in securing the conformity of his will to his second-order volitions, then, that a person exercises freedom of the will. And it is in the discrepancy between his will and his second-order volitions ... that a person ... does not have this freedom ...” (15)

**3. Applications:** Frankfurt’s proposal seems to get things right in a number of the cases that we have looked at. For instance:

- **Drug Addicts and Kleptomaniacs:** Imagine a drug addict who both wants a drug (because of the addiction) and doesn’t want it (because it is bad for her, or whatever). If she wishes she DIDN’T want the drug, but takes it anyway, she is not acting freely because she isn’t able to bring her will (i.e., her effective first-order desire) in line with her second-order volition (simply put, **she is not free to have the will that she wants to have**).

Note that, in theory, an addict might not have any second-order volitions about their addiction. They might not reflect or care one way or another which first-order desire wins out. Such an addict is merely a “wanton addict”. Such addicts have **freedom of action**—i.e., they want something and nothing stops them from DOING it—but they do not have **freedom of the WILL**. This is because they are not in control of WHICH desires move them to action.

*Note that there might also be “willing addicts”. Agents who WANT to be addicted. What Frankfurt says about this is dizzyingly confusing. I discuss it at the end.*

Similarly, kleptomaniacs who have an uncontrollable desire to steal, but wish that they could stop, are not in control of their actions. They are not acting freely. But, that seems intuitive. Both the unwilling addict and the unwilling kleptomaniac deserve HELP, not BLAME.

- **Jones on Election Day:** Recall Jones on Election Day. Even though Jones could not do otherwise, Jones still acts FREELY so long as the neurosurgeon does not intervene. Jones has a first-order desire to vote for Trump, and Jones IDENTIFIES with this desire, or reflectively APPROVES of it, so his vote is a free one.

However, if he wants to will to vote for Clinton, and the neurosurgeon intervenes to compel him to vote for Trump instead, he is NOT voting freely. For his will (i.e., his effective first-order desire to vote for Trump) is not in line with his volition (i.e., his second-order desire to will to vote for Clinton).

- **Animals:** Frankfurt's view nicely explains our intuition that animals do not have free will. Lacking second-order volitions, they are merely wantons. They are not blameworthy for their actions. If a lion mauls someone, we do not say "What a morally terrible lion! That lion had a moral obligation not to maul that person."

- **Toddlers, the Cognitively Impaired, Etc.:** Note that, for the same reasons that animals are not free, neither are human CHILDREN free (at least, toddlers aren't). This may seem counter-intuitive, but consider a 2-year-old child who pushes someone down some stairs. Are they morally blameworthy? Typically, we do not hold them responsible. Sometimes, we speak of the "**age of accountability**".

**On Freedom vs. Determinism:** In light of the Jones Election Day case, it is apparent that Frankfurt's view of free will is compatible with determinism. Whether or not you are able to do otherwise is irrelevant to free will. Consider Jones, when he votes WITHOUT be interfered with by the neurosurgeon. It is true of him that,

"he is not only free to do what he wants to do; he is also free to want what he wants to want. It seems to me that he has, in that case, all the freedom it is possible to desire or to conceive." (17)

**4. Objections:** Here are some problems with Frankfurt's proposal:

**(a) Whims:** We often "act on a whim". When we do so, we simply act on a (first-order) desire without reflecting on it or forming any second-order desires regarding it. But, then, this is a problem: If I find you annoying, and get a sudden urge to punch you in the face, and I do so, apparently I am not morally responsible for my action. "Hey!" I can say, "I just did it without reflecting on it first. I'm not responsible for that."

Reply #1: Perhaps one does not need to ACTUALLY engage in criteria (i) – (iv) in order to be morally responsible, but it just needs to be the case that one WOULD endorse their action (i.e., have a second-order desire to have the first-order desire).

Rebuttal: But, it's very likely the case that I WOULDN'T endorse my action of punching you in the face, had I taken the time to reflect on it. So, in that case, I STILL am not responsible for that action. (?)

*[We didn't cover this reply in class] Reply #2: In claim 1, we could remove the "only if", so that it instead says, "P is morally responsible for doing x if P does x freely." In philosophy-speak, the suggestion is that freedom is a SUFFICIENT condition for moral responsibility, but not a NECESSARY condition for it; i.e., maybe we could weaken Frankfurt's proposal to say that you're definitely responsible for actions that you perform freely, but maybe SOMETIMES you can be responsible even for actions that you do NOT perform freely.*

*Rebuttal: But, he doesn't say this. Also, this significantly weakens Frankfurt's view, and now we no longer have a full account of when one is morally responsible for one's actions and when one is not. The question would still remain unanswered, "So, then, what makes me responsible for my punching you!!?"*

Reply #3: When you act "on a whim" perhaps you don't question it because you DO identify with it. The only time that you would question a first-order desire is when it doesn't seem to cohere or "mesh well" with your other desires. When acting on a whim, there isn't any reason to engage in deliberation first. So, perhaps criterion (iii) is just automatically met in the case of acting on a whim. Frankfurt writes,

We "may suggest that volitions of the second order ... must be formed deliberately and that a person characteristically struggles to ensure that they are satisfied. But the conformity of a person's will to his higher-order volitions may be far more thoughtless and spontaneous than this. Some people are naturally moved by kindness when they want to be kind, and by nastiness when they want to be nasty, without any explicit forethought and without any need for energetic self-control. ... The enjoyment of freedom comes easily to some." (17)

In short, maybe the reflective endorsement can be either explicit OR implicit.  
[What do you think? Is this a good defense of Frankfurt?]

**(b) Weakness of the Will:** Perhaps the drug addict is not morally responsible for taking more of the drug if he wishes that he DIDN'T want the drug, but has a "weakness of will" where he cannot overcome the strength of that first-order desire. But, Frankfurt's thesis would entail that we are NEVER morally responsible for having weakness of will.

Imagine that you are tempted to cheat on your significant other. You do so. You reflect on the desire to cheat, and do NOT endorse it or identify with it. And yet, you find yourself doing it anyway. You suffered from "weakness of the will". On Frankfurt's view, you would not be morally blameworthy for such an action.

Reply: Disidentification with one's will is not MERELY thinking, "I shouldn't be doing this" or "this is not the right course of action." For Frankfurt, the struggle must be within the will itself. In other words, you must WILL the opposite of what you're doing. In the addict's case, the addict TRIES not to take the drug, or STRUGGLES against the urge, but fails. This is importantly different from the adulterer who thinks, "I know this is wrong, but what the hell. Here goes nothing!"

Rebuttal: Yet, surely some adulterers (or murderers, thieves, etc. fill in the blank here with any seemingly blameworthy action) actually TRY to resist the urge to act wrongly, and fail. So, Frankfurt's view still seems to entail that a pretty significant number of wrongdoers aren't blameworthy.

(c) **Manipulation:** Frankfurt claims to explain how one is free in manipulation cases such as Election Day. In that case, if the neurosurgeon were to press the button to control Jones's brain and force his vote, it would NOT count as free on Frankfurt's view because, even though he has a FIRST ORDER desire to vote for Trump, he has a second-order desire NOT to want to vote for Trump. Since Jones doesn't IDENTIFY with his action, his action is not free.

But, has Frankfurt just moved the problem back one level? What if the neurosurgeon's chip controls not only first-order desires but SECOND-order desires as well? In that case, Jones would be freely voting for Trump EVEN IF he is only doing so because the neurosurgeon has pushed the button to force him into doing so. That seems mistaken.

Reply #1: Perhaps we could appeal to THIRD-order desires?

Rebuttal: But, then we could just propose a case where the neurosurgeon manipulates those too. One could appeal, of course, to FOURTH-order desires... But, we are off on an infinite regress.

Reply #2: Perhaps we should expand the discussion to not merely include desire, but a whole range of things, including also values, beliefs, and so on. A free action would then be one that conforms to one's second-order desires, AND values, AND beliefs, etc.

Rebuttal: But, then we could just propose a manipulation case where the neurosurgeon manipulates THOSE TOO!

Reply: Perhaps we could add that, in addition to those things, it has to be the case that: If the agent knew the TRUTH about the SOURCE of those internal attitudes, s/he would STILL identify with them?

Rebuttal: But, we can't ever know these things for sure. So... We can't ever know for sure whether we are morally responsible for our actions?

**On Willing Addicts:** *Imagine that an addict reflects on her addiction and APPROVES of it (i.e., her will IS in line with her second-order volition). Confusingly, Frankfurt says that this "willing addict" is still NOT free. WTF? His reason seems to be that a criterion (v) is added. Something like, (v) A's will to X is (at least partially) CAUSED by her (second-order) reflection about which (effective first-order) desires she wants to have. He writes, "Suppose that a person has done what he wanted to do, that he did it because he wanted to do it, and that **the will by which he was moved when he did it was his will BECAUSE it was the will he wanted.**" (19)*

*If the addict's effective desire for the drug is not caused by any kind of second-order reflection or endorsement, but, rather, by the addiction, then she is not free. (But, if it IS partially caused by the second-order reflection, then it IS free??) Let's ignore this...*