

Free Will: Frankfurt & Watson

1. The “Consequence Argument” for the Incompatibilism of Determinism and Moral Responsibility

(*) Take a Person P, action A and time t: P A’s at t.

(**) Let P be born just after t-n.

(1) If the universe is deterministic, then the natural laws (e.g. the laws of physics) together with all the facts about the state of the universe at t-n, entail that P A’s at t.

Therefore,

(2) If the universe is deterministic, to do otherwise than A at t, P would have had to change the natural laws or alter facts about the universe at t-n.

(3) At no point in time could P have altered the natural laws or facts about the state of the universe at t-n. (Why? Because the laws are not under our control and t-n passed before P was born.)

Therefore,

(4) If the universe is deterministic, P could not have done otherwise than A at t.

(5) If P could not have done otherwise than A at t, then P did not A freely at t.

(6) If P did not A freely at t then either (a) P is not morally responsible for Aing at t, or
(b) P freely did something at some time prior to t that made it the case that she couldn’t do otherwise than A at t.

(7) Given that what determines that P A’s at t are the laws of nature and the state of the universe prior to when P was born, if determinism is true, then it is not the case that P did something to make it the case that she couldn’t but A at t.

Therefore,

(8) If the universe is deterministic, P is not morally responsible for Aing.

Therefore,

(9) If the universe is deterministic, no one is morally responsible for anything.

2. Frankfurtian Compatibilism

Frankfurt denies premise (5), premise (6) or both. He argues that one can lack the ability to do otherwise and still be morally responsible for what one does. His argument revolves around the case of a merely “counterfactual” intervener. S embezzles the money without the intervention of B. But if S had tried not to embezzle (if he had shown signs of not following through on his evil intention), B would have intervened and made S embezzle. (We’re to imagine that B has some way of exercising control over S when he wants to—a remotely controlled device in S’s brain perhaps.)

It isn’t clear whether Frankfurt thinks (a) S’s action is unfree but that S is still morally responsible for its performance, or whether (b) Frankfurt thinks S’s action is freely executed even though S couldn’t have acted otherwise. (Different articles of Frankfurt’s suggest different answers to this interpretive question. At one point he suggests that a person can act *from her own free will* even if *her will is not free*.)

“It is a mistake, however to believe that someone acts freely only when he is free to do whatever he wants or that he acts of his own free will only if his will is free” (1971, 19).

This seems like an outright contradiction to me, but maybe Frankfurt thinks “acted from his own free will” is idiomatic and that its correct application does not require that the agent who “acted from his own free will” was genuinely free to will one thing rather than another. In any event, taking the first option would have Frankfurt abandoning premise (6), the second would have him abandoning premise (5).

3. The Compatibilist’s Burden

Regardless, Frankfurt must shoulder the compatibilist’s burden. The Compatibilist must explain the difference between free action and unfree action and similarly distinguish action for which we’re morally responsible from action for which we’re not morally responsible. If freedom of action and/or moral responsibility doesn’t require the ability to do otherwise, what does it require?

4. Skepticism

The skeptic about unfreedom just identifies free action with intentional action. S As intentionally if the explanation for S’s acting is (or involves in the right way) the fact that S wanted to A. As Aristotle noted, intentional actions are explained by something *within* the agent (i.e. her motives). Unintentional behavior (e.g. being blown into someone by the wind or tripping into them or whatever) do not have these internal, psychological causes. If this is all there is to the distinction between mere behavior and manifestations of agency, **S acts freely when Aing if she then does what she wants to do when Aing.** It follows that every intentional action is a free action. If you did it intentionally, then you wanted to do it. If you wanted to do it, then you did it freely.

Question: Do I act freely when I hand over my wallet at gun point? Do I vote freely when I am threatened with death if I don’t vote for candidate X?

Hobbes says yes. You could have chosen to hold on to your wallet. So you did act freely. Still, the Hobbesian can say that your choices were limited by the thief in an immoral or unfair way. So you acted freely, but you were not free to do what you most wanted to do: i.e. keep your money and your life.

What if the desire that leads me to act as I do (or the psychological causes of my actions) are the products of brainwashing: the conscious manipulation of my motives and thinking by bad people who use drugs and torture to induce these motives in me? Does the fact that I act from these motives entail that I freely do what they’ve programmed me to do?

Science fiction is one thing, but consider the following example (taken from Wikipedia):

In 1991, the [Sierra Leone Civil War](#) started. Rebels invaded Ishmael Beah's hometown, Mogbwemo, located in the Southern Province of Sierra Leone, and he was forced to flee. Separated from his family, he spent months wandering south with a group of other boys. At the age of 13, he was forced to become a [child soldier](#). According to Beah's account, he fought for almost three years before being rescued by UNICEF. Beah fought for the government army against the rebels. In 1997, he fled [Freetown](#) by the help of the UNICEF due to the increasing violence and found his way to New York City, where he lived with Laura Simms, his foster mother. In New York City, Beah attended the [United Nations International School](#). After high school, he enrolled at [Oberlin College](#) and graduated in 2004 with a degree in [Political Science](#).

Beah says he doesn't remember how many people he killed during his time in the [Sierra Leonean government army](#). He and other soldiers smoked [marijuana](#) and sniffed [amphetamines](#) and

"brown-brown", a mix of cocaine and gunpowder. He blames the addictions and the brainwashing for his violence and cites them and the pressures of the army as reasons for his inability to escape on his own: "If you left, it was as good as being dead."

Questions: Did Beah freely kill those people he shot? Did he freely will these killings? Is he morally responsible for these killings?

5. Frankfurt distinguishes between freedom of action and freedom of will.

You have *freedom of action* if you're able to do what you want to do.

You have *freedom of will* if you're able to will what you want to will.

An agent's will is her set of "decisive" first-order desires: i.e. those desires that do (or would) move her all the way to action.

A desire to have a certain will, is what Frankfurt calls a *second-order volition*: it is a desire to have one of your first-order desires move you to action.

- (1) First-order desire: Your desire to become a rock star.
- (2) Second-order desire: Your desire to have (or keep) your desire to become a rock star.
- (3) Second-order volition: Your desire that your desire to become a rock star move you to act.

If you want to have your desire to become a rock star move you to act, then (according to Frankfurt) you *identify* with this desire. If you want to live a comfortable secure life, and you know that your desire to become a rock star may very well lead to actions that will risk your comfort and security (the music business is, after all, a risky career), identifying with your desire to become a rock star means wanting it to lead you to act – wanting it to win out over your desire for security in the "competition" between your desires over the determination of your behavior. If your fear of failure prevents you from pursuing the life of a rock star, but you still really identify with that desire (i.e. you still really want your desire to be a rock star to win out over your fear of failure), then your actions stem from a will that differs from the one you want.

Questions: If your will is not what you want it to be, mightn't you still have free will in Frankfurt's sense? Does this depend on whether you **could have had** the will you wanted to have? E.g., in the example above, in which your fears win out over your desire to be a rock star, we know you *didn't* will what you wanted to will. Is it compatible with that, that you **could've** willed what you wanted to will?

In any event, Frankfurt thinks **the satisfaction of first-order volitions is sufficient for possession of the kind of free will worth wanting:**

"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. Then he did it freely and of his own free will. Even supposing that he could have done otherwise, he would not have done otherwise: and even supposing that he could have had a different will, he would not have wanted his will to differ from what it was."

6. Watson's Criticisms of Frankfurt

1. What stops the infinite regress of higher-order desires?
2. What gives higher-order desires their *authority*?

Watson argues that we rarely form the second-order desires and volitions that are essential to Frankfurt's account of free will and moral responsibility. Instead, typical conflicts involve a tension between our **desires** and our **values**, where **our values are best correlated with our judgments about what we should do or what we have most reason to do**.

Typically, Watson says, our values are constituted by our desires so that we want to do what we judge we have most reason to do and when we judge that we should do something because it would be the best course of action, we therein come to want to do that thing because we judge it best. But according to Watson there are cases in which one can want something without valuing it at all.

Two Cases: (1) Feeling an urge to smother a baby who won't stop crying. (2) feeling an urge to smash an opponent in the face for being smug about his victory.

Question: Are these cases in which one wants (however momentarily) to smother the baby or smash the tennis player without valuing that action or outcome?

Illustrating the Difference between Frankfurt and Watson: Consider a third case that might help us distinguish these two views of moral responsibility in an extensional manner:

Mary is in a dead-end marriage. Her husband is verbally abusive and disrespectful. They have no children. Mary wants to get a divorce to improve her life. Moreover, Mary is an atheist, she does not believe in God, but because she was raised a religious catholic, she still thinks of divorce as (in her words) "a bad thing." Mary wants a divorce to rid herself of a constant source or abuse; Mary wants to remain married because she doesn't want to do something bad. Now suppose Mary reflects on these two desires, and comes to want that her desire to remain married have sway over her actions. That is, she wants her desire to stay married to "win out" for control over her actions. She wishes she didn't want a divorce. She is glad that she wants to stay married. But this isn't because Mary has arguments against divorce that she finds compelling. Her negative assessment of divorce remains entirely intuitional (or "system 1"). If Mary's desire to get divorced is nevertheless motivationally stronger than her desire to stay married, and she gets a divorce while feeling bad and disappointed in herself for ending her marriage in this way, does Mary act from a free will? Does she display autonomy?

Questions: What would Frankfurt say? What would Watson say? Which (if either) of them is right?

7. The Modality Issue for Frankfurt and Watson

Suppose we interpret Frankfurt as saying a person acts *from a free will* iff she acts in accordance with a second-order volition (i.e., she acts from a desire that she wants to have move her to act). How often does this happen?

Suppose we interpret Watson as saying that a person acts *from a free will* iff she acts from her values—i.e. she does the thing that she, in a cool reflective moment, believes she has the most reason to do. How often does this happen?

Question: Do these accounts lead to skepticism about the extent to which we exercise free will?

Suppose we interpret Frankfurt as saying a person is *morally responsible* for her action only if the desire that moved her to act was the one that she wanted to have move her to act. **Will this excuse people of responsibility for actions for which we (intuitively) think they should be**

held responsible? Can you avoid responsibility for your actions by demonstrating that you were conflicted at the time and (though you did what you wanted to do) you preferred acting on a different motive?

Suppose we interpret Watson as saying a person is morally responsible for her action only if she acts in accordance with her beliefs about what she has most reason to do. Will this excuse people of responsibility for actions for which we (intuitively) think they should be held responsible? **Can you avoid responsibility for your actions by demonstrating that you were conflicted at the time and (though you did what you wanted to do) you didn't act in accord with your values?**

Question: Do killers and rapists typically reflect on their desires and endorse them upon reflection as the desires they want to have move them to action? Do killers and rapists typically articulate views of the good life or the life most worth living and retain their desires to rape and murder as desires for activities they think constitutive of the good or choice-worthy life? If not, do Frankfurt's and Watson's accounts lead to substantive skepticism about the extent to which those who commit the most heinous crimes are morally responsible for their behaviors?

The (modal) solution: Say a person acts from a free will when Aing iff when she As she is **able** to refrain from acting in accordance with her second-order volition or value judgment regardless of whether or not she actually does so. Say a person is morally responsible for her action only if she **could have** acted in accordance with her second-order volitions regardless of whether she actually did so.

The "agent level ability" problem for the compatibilist: What does 'able' or 'could' mean here? Does it require that the agent have the ability to violate a law of nature? We're back where we started as far as the debate over the compatibility of free will and determinism.

8. Recklessness, Weakness and Compulsion

Suppose S is obese and recognizes the problems his obesity causes for his health and social life. S continues to eat chocolate.

Recklessness: S acts recklessly if S falsely believes that he has more reason to X than not X but S should have known better. (That is, S's ignorance of the fact that it would be better for him not to X is culpable or blameable.)

Recklessness is an impairment of judgment, not an impairment of will per se.

Weakness: S acts in a weak-willed (or akratic) fashion if S knows that he has more reason not to X than to X, S nevertheless wants to X more than not X, and S **could** have resisted his desire to X and so could have refrained from X-ing.

Compulsion: S acts in a compulsive manner if S knows that he has more reason not to X than to X, S nevertheless wants to X more than not X, and S **could not** have resisted his desire to X and so could not have refrained from X-ing.

Questions: What is it to resist a desire? What distinguishes someone who is weak willed from someone who is strong willed? What distinguishes someone who is weak willed from someone who is compulsive? Must an action be reckless or weak-willed to be blamable? If the universe is deterministic, does that imply that no one is reckless, weak-willed, or compulsive? What if the universe is not deterministic? Might we come to know that the universe is not deterministic in a

Kantian way, by recognizing that we can do what is right even when we want to do what is wrong?