

Alex Janiak

PHIL552: Metaphysics

04/19/2023

### Moral Responsibility without Metaphysical Freedom

The contemporary arguments for the incompatibility of Free Will with Determinism are undoubtedly persuasive. If these arguments are to be accepted, however, then either Indeterminism with Free Will must be accepted or we must reject Free Will. Furthermore, the arguments for the incompatibility of Free Will and Indeterminism are also difficult to refute (“The Mystery of Metaphysical Freedom”). That leaves us with a single alternative: accepting that we have no freedom to choose our actions – there is no Free Will. Without Free Will, we would be reduced to non-autonomous entities that have the ability for action and causation but lack the freedom to choose those actions. There are significant problems with this conclusion as well: it conflicts with our deep intuition that we possess Free Will. I, like Peter van Inwagen, believe that even philosophically-trained individuals who accept incompatibility still believe that, within their everyday lives, they can choose what they have for breakfast – barring any unusual circumstances (“The Mystery of Metaphysical Freedom”). This argument, however, is susceptible to the appeal to intuition fallacy: our intuitions can be influenced by factors that do not necessarily correspond to objective reality and therefore do not serve as a reliable metric for what is true. It is perfectly plausible that we do not have the ability to do otherwise but still intuit that we have Free Will because of some undiscovered evolutionary psychological mechanism that creates a false sense of choice, shielding us from existential crises. The specifics are not relevant – what matters is that our intuitions can sometimes lead us astray from the truth.

The second major – and more tangible – issue with accepting we lack Metaphysical Freedom is that Free Will appears necessary for Moral Responsibility. Moral Responsibility is the concept that agents can be held responsible for their decision based on some set of ethical standards or values. It includes attributing powers or capacities to that agent and viewing their behaviors as arising from these powers and capacities (Talbert). More formally, Moral Responsibility can be defined as follows:

- (1) An agent  $x$  is morally responsible for some action  $c$  at time  $t$  iff.  $x$  does  $c$  at  $t$ ,  $x$  has the ability to not do  $c$  at  $t$ , and  $c$  at  $t$  is considered morally right or wrong according to some ethical standard or system of values.

It should be noted that “does” encompasses both direct and indirect actions  $c$ . Clearly, without Free Will, this definition of Moral Responsibility becomes untenable because an agent could never have the ability not to do  $c$  at  $t$ . Consider the following scenario: You are speaking to a friend and for no apparent reason, they slap you. Understandably, you are infuriated and begin to explain that it is not morally acceptable to physically assault your friends without justification (assuming you believe unjustified physical aggression is wrong). Your friend responds by saying, “I had no other choice! I have no Free Will”. If you were to accept this definition of Moral Responsibility, you ought to concede and cease your anger. How could you blame them if they had no choice? Of course, most people believe that individuals can be blamed for their moral transgressions and would not accept such an excuse.

This forfeiture of Moral Responsibility entails that either your friend is lying and really had the choice not to slap you (i.e they have Free Will), in which case the arguments for the

incompatibility of Metaphysical Freedom with Determinism and Indeterminism are wrong, or our definition of Moral Responsibility is incorrect because your friend should still be morally responsible even if they did not have the freedom to not slap you.

It appears the argument has reached an impasse. Throughout the rest of this paper, I aim to break this deadlock by proposing an alternative formulation of moral responsibility that places the blame on an agent for their mental state or the collection of that agent's emotions, thoughts, and beliefs. This new, utilitarian definition allows agents to retain Moral Responsibility without committing to a stance of Free Will. Most importantly, this formulation makes a much more compelling case for the absence of Metaphysical Freedom by addressing its incompatibility with Moral Responsibility.

To understand why Moral Responsibility might not need to rely on Free Will, it's important to define accountability and attributability. Attributability is whether an action can be attributed to an individual:

- (2) An action  $c$  done at time  $t$  can be attributed to an agent  $x$  iff.  $x$  did  $c$  at  $t$  and  $x$  had the ability to not do  $c$  at  $t$ .

Attributability is the principle that assigns the cause or source of actions to individuals. Under the original definition of Moral Responsibility, without an action being attributed to an individual, that individual could not be held morally responsible for that action. In the previous example, your friend did not have the ability to not slap you when they did, so you couldn't attribute the slap to them. I.e, it was not their choice to slap you. Accountability, on the other hand, pertains to the degree to which individuals are held responsible for their actions:

- (3) An agent  $x$  is accountable for an action  $c$  iff.  $x$  is responsible for  $c$  and  $x$  is responsible for explaining, justifying, or accepting the consequences of  $c$ .

If an individual is morally responsible for an action, they are accountable for that action.

Your friend *did* slap you regardless of whether the slap was attributable to them, so it makes sense that they should be responsible for the slap and for apologizing. However, without Free Will, your friend still argues that they can not be held accountable because you can't attribute the choice to them.

The problem with your friend's new line of reasoning is that attributability and responsibility are distinct concepts, even if they can sometimes be used interchangeably. Oftentimes, a parent is assigned a degree of responsibility for their child's actions. If the child sets fire to a neighbor's house, the parent is held responsible for paying the damages. The arson might be attributed to the child, but the parent is still held accountable. The parent is held responsible for the child by society's judgment, even if they had no way of stopping the child at the time. Of course, one can argue that if the parents raised their child properly, the child would have never done such a thing. However, even if the parents provided overwhelming evidence that they had educated the child on the dangers and consequences of arson, the parents would still be held responsible to some degree because it is impossible that they did everything perfectly to control the child. From this, it is apparent that responsibility refers to the moral and social judgment that agents can be held responsible for a set of actions while attributability focuses on an agent being the cause or source of those actions (Talbert).

It would seem that in the case of your friend's unjustified physical aggression, they would still be held responsible by others, therefore they should still be held accountable for their actions regardless of whether they had the choice not to slap you because they are still responsible for the slap; they are responsible for their mental state and moral beliefs, and ultimately responsible for the decisions they make. This is the crux of the alternative formulation of Moral Responsibility: Moral Responsibility, rather than being associated with attributability, can coexist with a lack of Free Will as a form of accountability for an agent's mental state. This brings us to the proposed definition of Moral Responsibility without attributability:

- (4) An agent  $x$  is morally responsible for some action  $c$  at time  $t$  iff.  $x$  is accountable for  $c$  due to  $x$ 's set of mental states at  $t$ , and  $c$  is considered morally right or wrong according to some ethical standard or system of values.

When you hold an individual morally responsible for an action you are keeping their set of moral beliefs accountable by putting blame on the individual. You keep them accountable so that their beliefs and morals may change and in the future, their actions may be more moral. This definition of Moral Responsibility is utilitarian in nature and is aimed purely as a means of keeping an individual's set of moral beliefs in sync with that of society's values.

When applied to the previous example of your friend's physical aggression, you blame your friend for the slap even if they were not able to do otherwise, so that in the future, when they reach the decision whether to slap a friend, their moral beliefs are changed from the blame you had put on them and they are less likely to do so again. But how is your friend less likely to do some action if their decisions are predetermined under Determinism? Just because an agent

has no Free Will and can not do otherwise, doesn't necessarily mean that those actions can not be informed by their beliefs. Assuming the validity of the Consequence Argument, an agent's actions are causally necessitated by prior events and the laws of nature ("An Essay on Free Will"). Therefore, if an agent's set of moral beliefs reflects the accountability placed on them for a previous moral blunder, then the agent's future actions are a consequence of the natural laws and the past – which includes that accountability. Therefore, when an agent is held morally responsible for an action and their set of moral beliefs change, their future actions should reflect those different moral beliefs. When you give your friend advice on why it is morally unacceptable to hurt your friends, their future actions should reflect the belief that it's unacceptable to hurt your friends given they accept this belief.

One case that causes trouble for this proposed definition, as well as any other formulation of Moral Responsibility, is whether an agent can be held Morally Responsible if they are forced into a scenario in which they were not able to act on behalf of their moral beliefs. For example, perhaps your friend's brother is being held at gunpoint and their attackers said they have to slap you. In this case, your friend performs an immoral decision even if they were able to do otherwise. If they do not slap you, they cause their brother's death. If they slap you, they are hurting a friend. Is your friend still morally responsible for the slap in this new scenario? Deciding this is difficult and depends on various factors, the nature of the coercion, the choices available, and the moral principles involved – so it is beyond the scope of this paper. What can be said is that in a case such as this, the agent is treated as responsible for their mental state and still held accountable to the socially and morally judged degree depending on the situation.

Another similar potential case of confusion for this proposed definition of Moral Responsibility could be occurrences of Akrasia. Akrasia, defined originally by Aristotle and

Plato, is the phenomenon of acting against one's own better judgment or reason. With regards to Moral Responsibility, even if an agent has some moral belief, they may act against it. Suppose an individual is taught repeatedly that taking another person's life is morally wrong, and they have also explicitly stated that they believe that committing murder is immoral. Despite this, however, they still end up killing someone. In such cases of Akrasia, should agents be held morally responsible for their actions, even if the action does not align with their set of moral beliefs – specifically if they have no Free Will? Plato denied Akrasia and viewed it as an instance of ignorance or lack of understanding of what is truly good (Frede). Through the lens of this position, the person had either forgotten the belief that murder is immoral, never had the belief in the first place, or did not fully understand that murder is wrong and thought that *this* instance of murder was morally justified. As a result, the murderer is accountable for their faulty set of moral beliefs and is morally responsible for the murder.

On the other hand, Aristotle believed that Akrasia was an issue of self-control that arises when an individual's non-rational desires overwhelm their rational desires, leading them to act against their better judgment (Kraut). In this case, albeit the murderer might have had the moral belief that killing is wrong, they also had other desires and beliefs – like that the murder was justified – that were stronger than their moral beliefs. The murderer would thus be held responsible for their actions because you could hold their desires – a part of their mental state – accountable for conflicting with their moral beliefs. Ultimately, it is evident that under states of Akrasia – with both Plato's and Aristotle's interpretations – an agent can be held Morally Responsible even if they do not have Free Will because they are responsible for their mental state. It is important to mention that this argument assumes that people act solely according to their beliefs and mental state and thus can be held accountable for them.

Furthermore, this discussion can extend into more complex examples that question whether Akrasia undermines Moral Responsibility even with Free Will, for example when the agent is emotionally distressed or mentally ill. These cases are similar to the situation stated earlier when the agent is coerced by a third party and should be treated similarly. The proposed formulation of Moral Responsibility is *equivalent* to the original definition on a case by case basis except for these borderline cases where there is already confusion. That is, an agent is deemed morally responsible under the new definition if and only if the original definition finds that agent morally responsible in the same situation. This means that the new definition should encounter similar barriers of complexity along edge cases.

Although incompatibilists of both Determinism and Indeterminism have rather convincing arguments for the lack of Free Will, I find it really difficult to envision a reality without some sort of Moral Responsibility. Using the proposed definition of Moral Responsibility developed in this paper, agents can still be morally responsible for actions that they had no choice over because they are held accountable for their mental state. Additionally, under this definition, the purpose of keeping someone morally responsible for an action becomes utilitarian and simply seeks to keep society's moral values instilled in its members. This purely utilitarian purpose does mean that if everyone realized the true meaning of this Moral Responsibility, they could collectively come to an agreement that their actions can't really be attributed to them and reveal the façade. This might be a major flaw, but a similar problem arises when the individuals and the collective realize they don't have Free Will: their actions lose meaning. Even with this limitation, this proposed definition of Moral Responsibility makes accepting a lack of freedom much more compelling. Such a reality would mean that we do not have Free Will, but live as if we do.



## Works Cited

- Frede, Dorothea, and Mi-Kyoung Lee. "Plato's Ethics: An Overview." *Stanford Encyclopedia of Philosophy*, Stanford University, 1 Feb. 2023, <https://plato.stanford.edu/entries/plato-ethics/>.
- Kraut, Richard. "Aristotle's Ethics." *Stanford Encyclopedia of Philosophy*, Stanford University, 2 July 2022, <https://plato.stanford.edu/entries/aristotle-ethics/#Akra>.
- van Inwagen, Peter. "An Essay on Free Will." Oxford University Press, 1983.
- van Inwagen, Peter. "The Mystery of Metaphysical Freedom." *Metaphysics: The Big Questions*, Edited by Peter van Inwagen and Dean W. Zimmerman, Blackwell Publishers, 1998, pp. 365-374.
- Talbert, Matthew. "Moral Responsibility." *Stanford Encyclopedia of Philosophy*, Stanford University, 16 Oct. 2019, <https://plato.stanford.edu/entries/moral-responsibility/>.