

On Pereboom's Argument Against Compatibilism

Intro to Philosophy Paper #2

Ozaner Hansha

December 3, 2018

Abstract

There are 4 main goals of this paper. 1) Introduce the trichotomy of views that fall out of the irreconcilability of the free will thesis, determinism, and incompatibilism. 2) Extract Derk Pereboom's argument against compatibilism from his article "Why We Have No Free Will and Can Live Without It" [1]. 3) Elaborate on Pereboom's justification of said argument. 4) Analyze and critique this argument and its associated justification.

1 The Trichotomy of Free Will

1.1 The 3 Propositions

Before we delve into Pereboom's argument, we first need to clear up a couple of definitions:

- **The Free Will Thesis:** Some of the actions humans make are free. That is, while some of our actions may be forced or coerced in some way, there are certainly some actions that of our own accord.
- **Determinism:** Every event is caused by some prior event. That is, all events belong to an unbroken chain of events that would always have taken place given the past conditions.

- **Incompatibilism:** Determinism implies that no action we make is free (i.e. determinism implies the free will thesis is false) and vice versa. One cannot believe in both free will and determinism.

1.2 Justification

It would seem that, taken separately at least, all three of these propositions are reasonable:

- We generally believe ourselves as possessing the ability to make decisions of our own accord, that is we have **free will** (whatever we may define that as).
- Our intuitive understanding of cause and effect (as well as our training in classical physics) seem to imply that **determinism** is indeed the case.
- And finally, it would initially seem obvious that if all events are pre-determined, including our own thoughts and actions, then we truly could not have a choice to make. That is, free will and determinism are **incompatible** and we must give up one or the other.

Notice that the last proposition, incompatibilism, makes clear that these three statements cannot be true simultaneously. The only way to deal with this situation, then, is to reject at least one of these statements. Below we will briefly define the 3 different views that correspond to rejecting each statement.

1.3 The 3 Stances on Free Will

- **Hard Determinism:** Accept determinism and incompatibilism, and thus reject the free will thesis. This means that no one ever does anything freely.
- **Libertarianism:** Accept the free will thesis and incompatibilism, and thus reject determinism. This means that, at least for free

actions, some actions are not caused by previous events.

- **Soft Determinism:** Accept determinism and the free will thesis, and thus reject incompatibilism. This means that we can still act freely, despite all our actions having previous causes.

These are all the views that accept two of the statements and reject one. However, there are other stances that may reject two or even all three of the statements. The **Hard Incompatibilist** view that Pereboom will push is one such stance that simply argues against compatibilist views (views that reject incompatibilism).

2 Extraction

2.1 Structure of the Argument

The argument we are interested in takes place in section 2 of Pereboom's article, titled "Against Compatibilism". Pereboom's argument comes in 4 cases, all of which he designs to satisfy the common conditions of compatibilist free will (second order desire, able to reason rationally/morally, etc.). The hope is that the cases will get progressively more realistic with the last case being a legitimate challenge to the compatibilist of free will. At the same time, the argument depends on the idea that there is no real difference between any two consecutive cases that could allow one to create a new criterion for their preferred conditional definition of free will.

2.2 The 4 Cases

Case 1) Professor Plum is on the fence about killing Mrs. White. Even further he wants to want (has a second order desire) to do so. Moreover he can reason about the world rationally and understands morality. A team of scientists then send a radio wave to manipulate his brain and make his egoistic reasoning slightly stronger than his reason-

responsive thinking (yet not so much so that it is inconsistent with his standard self) thus pushing him to kill Mrs. White.

Case 2) The same as Case 1 but instead of the scientists manipulating his mind right before the event, they instead manipulate him at the beginning of his life, which casually determines him to make that decision to kill Mrs. White.

Case 3) The same as Case 2 except now, instead of scientists doing this to him, it was the way his household and community raised and trained him as a child. He was much too young to prevent resist this training.

Case 4) Plum is just a normal human like anyone else that satisfies the normal conditional definitions of free action, and decides to kill Mrs. White due to egoistic reasons.

2.3 The Argument

The argument can be formulated in the following way:

P1 Cases 1,2,3, and 4 satisfy all the common conditional definitions of free action.

P2 Professor Plum is not morally responsible in case 1.

P3 There is no difference in moral responsibility between cases 1 and 2

P4 There is no difference in moral responsibility between cases 2 and 3

P5 There is no difference in moral responsibility between cases 3 and 4

P6 Thus, Professor Plum is not morally responsible in case 4.

P7 However, if Plum is not morally responsible in case 4, nobody can be morally responsible.

P8 We have moral responsibility.

∴ Therefore, by the compatibilists' own definition of free action, we do not have free will (of a moral sort).

Pereboom phrases this as the “sort of free will required for being morally responsible”, and via this argument asserts that it cannot exist under the common conditional definition of free action.

He takes this further and says that it cannot exist under *any* reasonable conditional definition of free action as there is no moral difference between each case, meaning there is no distinguishing feature that makes Plum morally responsible in one and not the other.

3 Justification

3.1 Premise 1

As Pereboom made clear in his argument, all of his cases were constructed such that they satisfy all the most common compatibilist definitions of free action:

- They satisfy Hume’s “durable and constant character” [2] since Plum has always been willing to kill if the egoistic reasons outweigh the reason-responsive ones.
- They satisfy Frankfurt’s “second order desire” [3] since Plum wants to be the type of person who might kill for personal gain (this may be immoral but not irrational).
- They satisfy Fischer’s criterion that in some situations if the reasons are different the action wouldn’t have been performed [4], since if Plum knew he would face harsh consequences for killing Mrs. White he wouldn’t have done it.
- And finally, the cases satisfy Wallace’s criterion of the actor being able to grasp, reason, and regulate their behavior by “moral reasons” [5], since Plum is totally aware of the norms of morality and could stop himself if he felt morally necessary.

3.2 Premise 2

Pereboom justifies this via the following intuition:

“...intuitively, he is not morally responsible for the murder, because his action is causally determined by what the neuroscientists do, which is beyond his control.”

And this, he assumes, is intuitive enough to agree with any standard notion of moral responsibility. He considers the rebuttal that Plum may indeed not be acting freely due to this being a local manipulation of his brain state, but to that he counters:

“It is my sense that such a time lag, all by itself, would make no difference to whether an agent is responsible.”

Which again, is an intuitive enough proposition to accept. Why would free action be dependent on how long ago another agent coerced you act in some way?

3.3 Premises 3,4 & 5

3.3.1 Case 1 to 2

Pereboom’s justification of premise 3 is the same as his counter to the time rebuttal of premise 2:

“Again, it would seem unprincipled to claim that here, by contrast with Case 1, Plum is morally responsible because the length of time between the programming and the action...”

Thus reiterating that basing moral responsibility on an arbitrary time limit undermines it.

3.3.2 Case 2 to 3

Pereboom states that any challenge of this premise would have to show that Case 2 has some feature that morally distinguishes it from Case 3. He

says there is no such feature since in both cases Plums action was determined by factors far in the past that were out of control. Whether or not it was done by his community and household or a team of neuroscientists is irrelevant, he argues.

3.3.3 Case 3 to 4

Again, Pereboom does not see any significant distinguishing feature between cases 3 and 4. He notes one difference is that, unlike in case 3, case 4 was not caused by other agents (humans in this case) and instead by the chain of causality that envelopes all of our actions. He notes that this isn't morally distinguishing however because we can consider a case:

“...that is exactly the same as, say, Case 1 or Case 2, except that Plum’s states are induced by a spontaneously generated machine-a machine that has no intellegent designer.”

Here, Pereboom argues, Plum would still not be morally responsible.

3.4 Premise 6

It seems clear that if we accept premises 2-5 we must then accept premise 6. We can write the following where \leftrightarrow means the two cases are morally equivalent:

$$\text{Case 1} \leftrightarrow \text{Case2} \leftrightarrow \text{Case3} \leftrightarrow \text{Case4}$$

And since in Case 1 Plum is not morally responsible, he is not responsible in Case 4.

3.5 Premise 7

It is at this premise that Pereboom reaps the reward of setting up the first 3 cases. Indeed, this most general case encompasses all of human action, Pereboom concludes. This is because if physicalist determinism is true, then all Plum's (and anybody else's) actions would have been decided

beforehand regardless of how they were raised or if anyone manipulated them. And this is true regardless of whether they fit the conditions required of compatibilists for free action.

3.6 Premise 8

Premise 8 has been implied by Pereboom since the beginning of the argument. Presumably, he believes, that a compatibilist would agree with moral responsibility and its relation to free will. Thus he takes it as given to show that it is at odds with free will. He provides no other justification of it however.

4 Analysis

4.1 General Agreement

I would personally agree with Pereboom that moral responsibility and free will (at least the kind interesting kind strong enough to support moral responsibility) are at odds with each other and I think he does a good job at making explicit the hesitations some may have with compatibilism.

His argument strategy is effective in that it leads from a set of premises that a compatibilist would agree with (that in Case 1 Plum is not responsible and that we do in fact have moral responsibility), and over the course of 3 cases show that this case is totally analogous to the general case of determinism being incompatible with moral responsibility.

Moreover the actual implementation of the argument, which is mostly done via his cases, is also quite effective. I am hard pressed to find a morally distinguishing feature that a compatibilist would be able to roll into a conditional definition of free action. Pereboom addresses the obvious differences between each case and shows that they are not significant to our purposes (when agents cause it, which agents cause it, and agents

vs. non agents).

4.2 Regarding Moral Responsibility

That said, while I think Pereboom's argument is sound (insofar as I can see), I don't think it is valid. In particular, I object to premise 8, that we have moral responsibility. It would appear to me that there is no difference in positing the existence of free will vs. the existence of moral responsibility. Indeed most hard determinists, which Pereboom presumably is a form of since he rejects compatibilism and free will, believe that the rejection of moral responsibility (at least in its normal form) is required. And so, trying to show that compatibilist free will is false by assuming moral responsibility seems a bit troubling without proper justification.

Although to be fair, even though the entire article is about the existence of moral responsibility despite our lack of free will, the purpose of this particular argument was solely to show that compatibilism is inconsistent with our normal notions of moral responsibility and thus free will. And to that end, I think it accomplishes it.

5 Conclusion

References

- [1] Pereboom, Derk.
Living Without Free Will, Cambridge University Press, 2001
- [2] Hume, David.
Book Here, Publisher, XXXX
- [3] Frankfurt, Harry.
Book Here, Publisher, XXXX

- [4] Fischer, John.
Book Here, Publisher, XXXX
- [5] Wallace, Jay.
Book Here, Publisher, XXXX