

“I couldn’t have done otherwise”

The principle of alternate possibilities (PAP), according to which one is responsible only if one could have done otherwise, has long been at the centre of the philosophical debate about free will. How come that so many philosophers have considered this principle to be *intuitive*? The reason is possibly that, sometimes, when we have done something wrong, we offer the following excuses: “I couldn’t have done otherwise”, “I had no other choice”. So, when philosophers say that the PAP is intuitive, the best way to understand this claim is to say that the PAP is somewhat embedded in our daily practices of excuses, rather than to say that it is self-evident.

If this is true (and we’ll try to find out whether it is), then the best way to understand the PAP is to understand what we mean by “I couldn’t have done otherwise” when we give such excuses. A good interpretation of the PAP cannot be at odds with our use of such excuses. The problem is that I have gathered empirical evidence that traditional interpretations of the PAP cannot account for such excuses. Consider for example the following case:

Because I had to travel, I borrowed my best friend’s car. As I am driving through a little town, a child suddenly crosses the road, chasing a ball. Though I’m moving at regular speed, the only way not to hit the child and kill him is to crash my friend’s car on a near way wall. I decide to crash the car to save the child. When my friend learns what happened to his car, he gets mad. Looking for excuses, I say: “I’m sorry. But I couldn’t have done otherwise. I had no choice.”

In this case, most subjects agree to say that I’m not to blame for having destroyed the car and I couldn’t have done otherwise.

Let’s first examine a compatibilist interpretation according to which “I could have done X” means “If I had decided to do X, I would have done X”. In this case, had I decided to save the car rather than the child, I would have hit the child. So, according to this interpretation, I could have done otherwise, and the compatibilist interpretation fails. All the same, there is no compatibilist interpretation of the PAP that would conclude that, in this case, I couldn’t have done otherwise. Clearly, someone who would choose to hit the boy and run away would be a total jerk but he would have total control upon his action – the kind of control required by incompatibilists. So, both compatibilist and incompatibilist interpretations of the PAP fail to account for our use of “couldn’t do otherwise”, which means that these interpretations of the PAP are not really *intuitive* and should not be accepted in absence of other reasons.

To account for these data, I’ll propose a new semantic for the use of “could have done otherwise” for an excuse that distinguishes between responsibility for bad events and responsibility for good events. Here is my proposition for responsibility for bad events:

An agent who performed a bad action A “could have done otherwise” if, had the agent had the preferences of a reasonably moral agent, he wouldn’t have done A.

And for good events:

An agent who performed a good action A “could have done otherwise” if, had the agent had preferences worse than the preferences of a reasonably moral agent, he wouldn’t have done A.

Where the context sets “how much worse” preferences we must consider.

Let’s now return to my crashing my friend’s car to spare the child. In this case, I think that we would expect of any decent man that he would prefer the child to the car. Clearly, we would think of someone giving more importance to the car as a real jerk. So, in this case, if I had the minimally good preferences, the preferences of the average man, I would still have crashed the car. Consequently, “I couldn’t have done otherwise”.

Finally, I’ll argue that this semantic can account for the results obtained by Phillips and Knobe about our use of the expression “being forced to”. Going further, I’ll present supplementary empirical data that fits our theory but directly go against Philipp and Knobe’s theory. In these scenarios, agents have to choose between two actions A and B, where A is good and B bad, while they believe that A is bad and B good. Finally, they choose B. Philipp and Knobe’s account predicts that these agents “were not forced to choose B”, since they do the bad thing. Our account predicts that these agents “were force to choose B”, since they do what they think to be the right thing to do. Our results show that subjects indeed consider people choosing B as force to choose B.