

A challenge for Frankfurt-style compatibilists

Philip Swenson

Published online: 1 July 2014
© Springer Science+Business Media Dordrecht 2014

Abstract The principle of alternative possibilities (PAP) tells us that an agent is morally responsible for an action only if he could have done otherwise. Frankfurt-style cases (FSCs) provide an extremely influential challenge to the PAP (Frankfurt, *J Philos* 66:829–839, 1969). And Frankfurt-style compatibilists are motivated to accept compatibilism about responsibility and determinism in part due to FSCs. But there is a significant tension between our judgments about responsibility in FSCs and our judgments about responsibility in certain omissions cases. This tension has thus far largely been treated as an internal puzzle for defenders of FSCs to solve. My goal here is to regiment this tension into a clear argument which (if sound) undermines the FSC based critique of PAP. I will also argue that there is an in principle reason to doubt that Frankfurt-Style Compatibilists will be able to successfully respond to my argument.

Keywords Moral responsibility · Frankfurt-style cases · Frankfurt-style compatibilism

Here is a FSC quite similar to the one originally presented by Frankfurt:

Original Frankfurt Case: Black wishes Jones to cast his vote for presidential candidate A. In order to ensure that Jones does this, he implants a chip in Jones's brain which allows him to control Jones's behavior in the voting booth. (Jones has no idea about any of this.) Black prefers that Jones vote for candidate A on his own. But if Jones starts to become inclined to vote for anyone other than A, Black will immediately use his chip to cause Jones to

P. Swenson (✉)
Department of Philosophy, University of California, Riverside, HMNSS Building, Room 1604, 900
University Avenue, Riverside, CA 92521, USA
e-mail: pswen001@gmail.com

vote for candidate A instead. As it turns out, though, Jones votes for candidate A on his own and Black never exerts any causal influence on Jones's behavior.

Initially it would seem that both of the following are true. (1) Jones is morally responsible for voting for candidate A. And (2) Jones could not have done otherwise than he in fact did. Thus we have an apparent counterexample to PAP. But I will argue that we have significant reason to doubt whether Jones is responsible.¹ My strategy will be to begin with cases in which it is clear that an agent is not responsible for his behavior (or lack thereof) and argue that our judgment concerning responsibility in these cases ought to be the same as our judgment in the Frankfurt-style cases.

Consider the following case offered by Fischer and Ravizza:

Sharks: John is walking along the beach and sees a child drowning in the water. John believes that he could rescue the child without much effort. Due to his laziness, he decides not to attempt to rescue the child. The child drowns. Unbeknownst to John, there is a school of sharks hidden beneath the water. If John had attempted to rescue the child, the sharks would have eaten him and his rescue attempt would have been unsuccessful.²

Note the similarity between the role of the sharks in this case and the role of Black in Original Frankfurt Case. Both the sharks and Black alter what the agent can do by playing the role of counterfactual interveners. Neither of them plays an actual sequence role in producing the agent's action or omission. But our judgments about how the sharks impact the agent's responsibility differ from our judgments about Black.

In Sharks John is responsible for something quite important. He is responsible for *not trying* to save the child. But, importantly for our purposes, it also seems clear that John is not responsible for *failing to save the child*.³ Of course, if the sharks had not been present, John would have been responsible for failing to save the child. So the sharks' presence seems to alter what John is responsible for while Black's presence does not seem to alter what Jones is responsible for. These disparate reactions provide the basis for my argument. In Sect. 1 I will lay out my argument and in Sect. 2 I will argue that there is an in-principle reason to doubt that Frankfurt-Style Compatibilists will be able to successfully respond to my argument.

1 The Challenge

This claim will function as the first premise of my argument:

¹ Here and throughout it should be assumed that by 'responsible' I mean morally responsible (unless I explicitly state otherwise).

² Fischer and Ravizza (1998, p. 125).

³ I know of no one in the literature who has claimed that he is.

(P1) In Sharks John is not responsible for failing to save the child.

Now consider a modified version of Sharks:

Penned-in Sharks: Everything occurs just as in Sharks except for the fact that the sharks are penned up. However, unbeknownst to John, there is an evil observer who wishes for the child to drown. If John had jumped into the water, the evil observer would have released the sharks, and as a result, the sharks would still have prevented John from rescuing the child. But the presence of the observer plays no role in the actual sequence of events.⁴

It seems to me that we ought to say the same thing about John's responsibility in Penned-in Sharks as we say in Sharks. Thus:

(P2) If John is not responsible for failing to save the child in Sharks, then he is not responsible for failing to save the child in Penned-in Sharks.

Initially it is hard to see why the mere fact that the sharks are penned up but would be released if John enters the water could make a difference to John's responsibility. Thus, I take it that anyone who would claim that (P2) is false would have some explaining to do. They would need to identify a moral responsibility-relevant difference between Sharks and Penned-in Sharks.⁵ Moving on for the meantime, let us now consider a third case:

Sloth: In this case, there are no sharks present to prevent a rescue by John. The evil observer is now monitoring John's thoughts instead. John decides (without deliberating much) to refrain from saving the child. If John had seriously considered attempting to rescue the child, the evil observer would have caused him to experience an irresistible urge to refrain from saving the child. However, this observer still plays no role in causing John's decision to refrain from attempting a rescue.⁶

We have now come to a case where it does seem intuitive that John is responsible for failing to save the child. Despite this, my view is that John's responsibility in Sloth should not differ from his responsibility in Penned-in Sharks. Thus, I suggest:

(P3) If John is not responsible for failing to save the child in Penned in Sharks, then he is not responsible for failing to save the child in Sloth.

The difference between Penned-in Sharks and Sloth is that in Penned-in Sharks the counterfactual intervention would have occurred after John's decision to try to help, while in Sloth the counterfactual intervention would have occurred before John

⁴ This case is also from Fischer and Ravizza (1998, p. 138). They credit David Kaplan for suggesting the case.

⁵ Fischer and Ravizza (1998) and Byrd (2007, pp. 56–67) attempt to do this.

⁶ This sort of case was suggested by Frankfurt (1994, pp. 620–623).

makes a decision. It seems strange to hold that a mere difference in the timing of the counterfactual intervention would result in a difference concerning John's responsibility.⁷

Let us now consider one last case:

Hero: John decides (without deliberating much) to rescue the child, and he successfully does so. Unbeknownst to him, if he had seriously considered refrain from rescuing the child, our now benevolent observer would have caused him to immediately experience an irresistible urge to rescue the child.⁸

Again I see no reason to think that John's responsibility should differ from Sloth to Hero. So I endorse:

(P4) If John is not responsible for failing to save the child in Sloth, then he is not responsible for saving the child in Hero.

I suspect that most will join me in finding (P4) to be intuitively obvious. One way to reject (P4) would be to argue, as Fischer and Ravizza once did, that the ability to do otherwise is required in order to be responsible for an omission, but not in order to be responsible for an action (call this view *The Asymmetry Thesis*).⁹ Harry Frankfurt introduced the Sloth case in order to show that we should not accept *The Asymmetry Thesis*, and it seems clear to me that he is correct about this.¹⁰ (Fischer and Ravizza later came to agree with Frankfurt on this point.¹¹)

We can now present my central argument in its entirety. (Let's call it the "No Principled Difference Argument"):

(P1) In Sharks John is not responsible for failing to save the child.

(P2) If John is not responsible for failing to save the child in Sharks, then he is not responsible for failing to save the child in Panned-in Sharks.

(P3) If John is not responsible for failing to save the child in Panned-in Sharks, then he is not responsible for failing to save the child in Sloth.

(P4) If John is not responsible for failing to save the child in Sloth, then he is not responsible for saving the child in Hero.

Thus;

(Conclusion) John is not responsible for saving the child in Hero.

Now since Hero is structurally identical to Original Frankfurt Case, the cogency of this argument gives us reason to doubt that Jones is responsible for voting as he does. And of course Original Frankfurt Case is only a counterexample to PAP if

⁷ Fischer and Ravizza (1998, pp. 140–141) make this critical point. Randolph Clarke holds the view that the timing of the counterfactual intervention does matter. See Clarke (1994, pp. 195–208) and Clarke (2011, pp. 594–624).

⁸ This case is drawn from Fischer and Ravizza (1991, pp. 258–278).

⁹ Fischer and Ravizza (1991). Sartorio (2005) argues that Sloth is not a counterexample to *The Asymmetry Thesis*.

¹⁰ Frankfurt (1994).

¹¹ See Fischer and Ravizza (1998).

Jones is in fact responsible in Original Frankfurt Case. So the challenge to those who wish to wield FSCs in a critique of PAP is to show that we can draw a principled line somewhere between Sharks and Hero such that we can plausibly maintain that John is responsible in Hero but not in Sharks. Accomplishing this would of course require rejecting (P2), (P3) or (P4). I will now argue that that there is good reason to think that this challenge cannot be met.

2 Why the Challenge is a Difficult One

One might think that the no principled difference argument poses no serious threat to defenders of FSCs. All they need to do to respond is just find a place to draw a line somewhere between Sharks and Hero. Indeed several accounts in the literature attempt to do this.¹² Rather than engage with each of these individual accounts in this short paper, I now want to provide a general reason for thinking that one cannot plausibly respond to the No Principled Difference Argument and continue to endorse the FSC based critique of PAP.

The paper in which Frankfurt first introduced FSCs contains the following remarks on the unique feature of FSCs in virtue of which they appear to be counterexamples to PAP:

“There may be circumstances that constitute sufficient conditions for a certain action to be performed by someone and that therefore make it impossible for the person to do otherwise, but that do not actually impel the person or in any way produce his action...An examination of situations characterized by circumstances of this sort cast doubt, I believe, on the relevance to questions of moral responsibility of the fact that a person who has done something could not have done otherwise.”¹³

In a similar vein, John Martin Fischer has pointed us to an important distinction between the “A-Factors” of a situation, which bring about a particular event, and the “B-Factors” which render the event inevitable but need not cause or bring about the event.¹⁴ Combining Fischer’s terminology with Frankfurt’s suggestion yields the following account. The reason it is so intuitively plausible that responsibility is not undermined in the FSCs is that FSCs are purportedly cases in which the factors that remove the agent’s ability to do otherwise are mere B-Factors,¹⁵ and mere B-Factors appear to be irrelevant to whether an agent has the type of control he needs in order to be moral responsible. The presence of Black in Original Frankfurt Case, for

¹² See Fischer and Ravizza (1998), Sartorio (2005) and Clarke (2011). I criticize each of these individual accounts in “Omissions and the Frankfurt cases: A challenge” (Unpublished Manuscript).

¹³ Frankfurt (1969).

¹⁴ Fischer (2010, pp. 267–278). See p. 269.

¹⁵ ‘mere’ in the sense that they are not also A-Factors.

example, appears to make it impossible for Jones to avoid voting for candidate A without actually causing Jones to vote for A.¹⁶

I think that Frankfurt has correctly identified the feature of the FSCs in virtue of which they appear to be counterexamples to PAP. To see this, imagine variations on the FSCs in which the intervener does play an active causal role in bringing about the agent's act (and is thus not a mere B-Factor). In these variants it becomes much less intuitive to hold that the agent's responsibility is unaffected by the presence of the intervener. The reason Black appears to be irrelevant in Original Frankfurt Case is precisely because he does not make anything happen in the actual sequence of events. Thus it appears that the principle underlying our intuitions about FSCs is that *mere B-Factors are irrelevant to moral responsibility*.

This reveals an additional challenge for anyone who wishes to reply to the no principled difference argument. The problem is that each case appealed to in the no principled difference argument centrally involves the presence of a mere B-Factor (the sharks in Sharks, the evil observer in Penned-in Sharks, etc.). Furthermore, accepting that the agent is not responsible in any of the cases apparently involves rejecting the claim that mere B-Factors are always irrelevant to moral responsibility. Frankfurt-Style Compatibilists (and any other defenders of FSCs) should say that the principle underlying our intuitions about FSCs is correct. So they should not accept the claim that John is not responsible in any of the cases appealed to in the no principled difference argument. Thus, they cannot plausibly draw a line anywhere between Sharks and Hero with regard to John's responsibility.

There are several potential lines of reply to this conclusion. One would be to claim that the presence of the sharks (in Sharks) does play a causal role in the actual sequence and thus is not a mere B-Factor. The difficulty with this reply is that it is quite counterintuitive to hold that the sharks cause you to fail to save the child.

A second line of reply would be to reject the claim that it is because Black is a mere B-Factor that it seems to us that he is irrelevant to moral responsibility. How plausible this reply would be depends to some degree on the details of the rival account, but it does seem that there is a significant cost to rejecting this claim. Again, it seems that Frankfurt was right in suggesting that Black is irrelevant precisely because he does not make anything happen in the actual sequence of events.

Thus it appears that replying to the no principled difference argument involves rejecting the intuitive principle underlying FSCs. Frankfurt-style compatibilists (and any other defenders of FSCs) should say that the principle underlying our intuitions about FSCs is correct. So the no principled difference argument is a serious problem for Frankfurt-style compatibilists. Cases like Sharks do not merely generate an internal puzzle for defenders of FSCs to solve. They threaten to undermine Frankfurt-style compatibilism.

¹⁶ Given this picture, one way to understand the dilemma defense is that it calls into question whether it is true that a mere B-Factor can rule out the ability to do otherwise.

Acknowledgments For helpful comments on this paper thanks to Randolph Clarke, D. Justin Coates, Chris Franklin, Joshua Hollowell, Ben Mitchell-Yellin, Garrett Pendergraft, John Perry, Michael Nelson, Adam R Thompson and Patrick Todd. Thanks also to the UCR Agency Writing Workshop and to audiences at the Central APA and the Tennessee Value and Agency Workshop. And extra-special thanks are due to John Martin Fischer for his significant help and encouragement.

References

- Byrd, J. (2007). Moral responsibility and omissions. *Philosophical Quarterly*, 57, 56–67.
- Clarke, R. (1994). Ability and responsibility for omissions. *Philosophical Studies*, 73, 195–208.
- Clarke, R. (2011). Omissions, responsibility, and symmetry. *Philosophy and Phenomenological Research*, 82(3), 594–624.
- Fischer, J. M. (2010). Replies. *Philosophy and Phenomenological Research*, LXXX(1), 267–278.
- Fischer, J. M., & Ravizza, M. (1991). Responsibility and inevitability. *Ethics*, 101, 258–278.
- Fischer, J. M., & Ravizza, M. (1998). *Responsibility and control: a theory of moral responsibility*. Cambridge: Cambridge University Press.
- Frankfurt, H. (1969). Alternate possibilities and moral responsibility. *Journal of Philosophy*, 66, 829–839.
- Frankfurt, H. (1994). An alleged asymmetry between actions and omissions. *Ethics*, 104, 620–623.
- Sartorio, C. (2005). A new asymmetry between actions and omissions. *Noûs*, 39, 460–482.