

MORAL RESPONSIBILITY AND THE PRINCIPLE OF AVOIDABLE BLAME

Gerald K. Harrison

University of Durham

***Abstract:** Many now accept that Frankfurt-style cases refute the principle of alternative possibilities (PAP). But, in this paper I argue that even if Frankfurt-style cases refute PAP they do not refute a related principle: the principle of avoidable blame (PAB). My argument develops from the observation that an agent in a Frankfurt-style case can be aware of the nature of their situation without this undermining their moral responsibility. I then argue that PAB captures all that is important about PAP such that the refutation of PAP is of no concern as long as PAB survives.*

***Key words:** moral responsibility, blame, Frankfurt-style cases.*

According to the principle of alternative possibilities one is morally responsible for what one has done only if one could have done otherwise.

PAP: A person is morally responsible for what he has done only if he could have done otherwise. (Frankfurt 1969)

Many consider the principle (hereafter referred to as PAP) to be refuted by a particular kind of counterexample, known as a Frankfurt-style case. In what follows I will not dispute that Frankfurt-style cases refute PAP. But I will contend that another principle, the principle of avoidable blame, is not refuted by Frankfurt-style cases. I will then argue that all that was important about PAP is captured by the principle of avoidable blame.

I will start by outlining Frankfurt's famous counterexample to PAP. I will then go on to argue that the principle of avoidable blame is not necessarily refuted just because PAP is. This by itself will not show that Frankfurt-style cases do not refute the principle of avoidable blame. Indeed, Fischer has recently argued that both principles are refuted by such counterexamples. I will consider Fischer's arguments and find that they fail. Finally I will suggest that the only reason we ever cared about PAP is for reasons to do with the principle of avoidable blame, and therefore so long as the latter principle is not refuted we should be unconcerned that PAP is false.

1. Frankfurt's refutation of PAP

PAP is an intuitive principle with a respectable history. Aristotle, for instance, invokes PAP when he says that "when acting is up to us, so is not acting" (1915: 1113b6). And Kant too invokes PAP when he claims that freedom entails that "the act as well as its opposite must be

within the power of the subject at the moment of its taking place” (1960: 45). And indeed, the principle enjoyed almost universal acceptance until in 1969 Frankfurt developed a certain thought experiment designed to call PAP into question.¹

Suppose someone – Black, let us say – wants Jones to perform a certain action. Black is prepared to go to considerable lengths to get his way, but he prefers to avoid showing his hand unnecessarily. So he waits until Jones is about to make up his mind what to do, and he does nothing unless it is clear to him (Black is an excellent judge of such things) that Jones is going to decide to do something other than what he wants him to do. If it does become clear that Jones is going to decide to do something else, Black takes effective steps to ensure that Jones decides to do, and that he does do, what he wants him to do. Whatever Jones’s initial preferences and inclinations, then, Black will have his way ...

Now suppose that Black never has to show his hand because Jones, for reasons of his own, decides to perform and does perform the very action Black wants him to perform. In that case, it seems clear, Jones will bear precisely the same moral responsibility for what he does as he would have borne if Black had not been ready to take steps to ensure that he do it. (Frankfurt 1969, pp. 835-36)

The defining feature of a Frankfurt-style case is that some person, or machine, or natural mechanism, does not intervene in a process leading to an agent’s action, but would have intervened if the agent had been about to decide to perform some different action. What such cases draw attention to is the fact that, given the counterfactual intervener, or machine, or natural mechanism, plays no role in the agent’s deliberations and subsequent action, it seems intuitive to consider that the agent is morally responsible for their action. And, given the presence of the counterfactual intervener (or machine, or natural mechanism) it also seems true to say that the agent could not have done otherwise. So, PAP is refuted because in a Frankfurt-style case an agent is morally responsible, yet they also could not have done otherwise.

Since Frankfurt first published his article, there have been those who have disputed that Frankfurt-style cases are successful counterexamples to PAP.² This debate continues, however it is a debate that I will ignore here. I will simply declare an allegiance to those who believe Frankfurt-style cases successfully refute PAP. In what follows then, I will simply take it for granted that in a Frankfurt-style case that an agent is blameworthy, yet could not have done otherwise than what they did do.

Rather, I believe that what motivates a commitment to PAP is the intuition that an agent needs to have the ability to avoid blame – they need, in other words, control over keeping their moral slates clean. And I contend that in a Frankfurt-style case the agent retains this ability – the ability to ensure that their moral slate is kept clean – even if they lack completely the ability to do

otherwise.

There are others who have made defences of versions of PAB. I want to distance myself from some of these versions of PAB however. Wyma, for instance, has defended what he has called the “principle of possibly passing the buck” (PPPB).

PPPB: A person is morally responsible for something she has done, A, only if she has failed to do something she could have done, B, such that doing B would have rendered her morally non-responsible for A. (1997, p. 59)

In a similar vein, Otsuka’s version of PAB states that:

OPAB: One is blameworthy for performing an act of a given type only if one could instead have behaved in a manner for which one would have been entirely blameless. (1998, p. 688)

Let me make clear that I believe that both PPPB and OPAB are false. They are not yet principles that can be defended against Frankfurt-style cases. In a Frankfurt-style case an agent is morally blameworthy, yet did not have the ability to do anything else. Both Otsuka and Wyma dispute that the agent in a Frankfurt-style case lacked the ability to do anything else. So, Otsuka and Wyma believe that some version of the principle of alternate possibilities needs to be true if either PPPB or OPAB are to be true. This is exactly what I deny. It is no condition of blameworthiness then, that the agent have the ability to *do something* that, in doing, would render them blameless. The following from McKenna is closer to the position I am defending:

As McKenna puts it, we “hold people accountable for what they do only if they can avoid any blame or punishment that might fall upon them for performing those very particular actions which they do perform” (1997, p. 74). The point then, is that it is a condition of being blameworthy that one could have avoided incurring blame. This condition is compatible with being unable to do otherwise, or so I claim. In a Frankfurt-style case Jones may not have been able to ensure that he make some alternative decision, but he could ensure that he would not be blameworthy.

To make my case I want first to draw attention to a common, and hitherto unquestioned assumption made regarding Frankfurt-style cases, and it is this: that the agent concerned is unaware of the precise nature of their situation. As Alison McIntyre says:

We view the agents in these Frankfurt-style examples with dramatic irony: they think that they could have chosen otherwise and done otherwise, but their belief is false. They are deluded in believing that they possess a kind of control that they actually lack. (1994, p. 455)

But we should note that it is an inessential feature of a Frankfurt-style case that the agent in is unaware of the presence of the counterfactual intervener. Consider that in his original article Frankfurt was concerned to argue that:

A person may do something in circumstances that leave him no alternative to doing it, without these circumstances actually moving him or leading him to do it – without them playing any role, indeed, in bringing it about that he does what he does. (1969, p. 830)

The above is consistent with the agent being aware of the situation that they are in. Jones might believe that he will be made to decide to assassinate the president unless he freely decides himself to assassinate the president, yet this belief play no role in his coming to the decision to assassinate the president. In this case, Jones's knowledge of his situation would not affect our assessment of his blameworthiness. Jones is well aware that he cannot do otherwise than decide to assassinate the president, yet it remains intuitive to consider him blameworthy for doing so where he chooses to do so on his own. In other words, the common assumption that the agent needs to be unaware of the Frankfurt-style nature of their situation is false.

In fact, we can go one step further and even allow that Jones's belief that he is in a Frankfurt-type case *does* play a role in his decision without this affecting our judgement of his moral responsibility. After all, if Jones freely makes the decision to assassinate the president, why should he not be blameworthy for doing so, even where a belief that he was in a Frankfurt-type case played a role? For instance, Jones might reason that because it is inevitable that he will decide to assassinate the president, he might as well freely decide on his own. We can imagine him reasoning as follows: "if I do not decide to assassinate the president I am going to be made to, so what the heck, I'll decide to do it myself" and freely decide to take the shot. In this case we will still consider Jones blameworthy. Or Jones might reason that because he cannot decide otherwise he cannot be held responsible for his decision, and therefore conclude that he has now an opportunity to freely and non-culpably decide to actually kill someone. If Jones reasons in this way and freely decides to shoot the president then it remains intuitive to consider him fully blameworthy, even where we are aware that Jones would not have decided to shoot the president had it not been for the Frankfurt-style nature of his situation. I will not press these points further, for it suffices for my argument that we acknowledge my weaker claim that Jones can be aware that he will be made to decide to assassinate the president if he does not decide to do so his own, and this belief play no role in his making the decision to assassinate the president.

Once we acknowledge that Jones can be fully aware that he is in a Frankfurt-style situation without this undermining his blameworthiness then it is also evident that in such Frankfurt-style cases Jones has the ability to avoid blame. Jones knows that he is going to make the decision

to assassinate the president, but he also knows that it is up to him whether he makes this decision himself, or is forced to make it. Jones, it is true, cannot do anything to avoid making the decision. But he still has the ability to avoid blame. He knows it is inevitable that he will decide to assassinate the president, but he also knows that it is not inevitable that he will be blameworthy for doing so. The latter matter is fully under Jones's control.

This does, I believe, show that the ability to avoid blame does not require the ability to do otherwise. Here I rely simply upon what it seems natural to say about Jones's abilities in such a situation, and do not propose to offer some analysis of the concept of ability. In other words, rather than offer an analysis I take it instead that it is a test of an analysis of ability that it delivers the right verdict regarding Jones. However, even if I am right and the ability to avoid blame does not require the ability to do otherwise, it might be objected that nothing I have said so far shows that Frankfurt-style cases do not refute PAB as well as PAP. To do this, I would have to show that making Jones ignorant of the fact that he is in a Frankfurt-type case would not have the effect of robbing Jones of the ability to avoid blame. Consider that so far I have argued that knowledge of the Frankfurt-style nature of his situation does not affect Jones's blameworthiness. But if PAB survives only where Jones knows he is in a Frankfurt-style situation, then PAB is still undermined by Frankfurt-style cases in which Jones is ignorant of the Frankfurt-style nature of his situation.

Why might removing Jones's knowledge of the Frankfurt-style nature of his situation mean that he no longer has the ability to avoid blame? In other words, why might it be claimed that a Frankfurt-style case refutes both PAP and PAB? John Martin Fischer argues that both PAP and PAB are refuted on the following grounds:

Now it might be thought that the possibility of avoiding authorship or the possibility of avoiding moral responsibility would be a more substantial basis for moral responsibility. But I believe there ... problems here. Note that in the alternative sequence in a Frankfurt-type case the agent would indeed be avoiding (say) moral responsibility, but he would be doing so "accidentally." The agent would not be voluntarily avoiding responsibility. The suggestion that avoiding responsibility is a sufficiently robust basis for moral responsibility may get some of its plausibility from the fact that in a typical context in which we would say that someone has avoided (say) blameworthiness, it would be in virtue of some voluntary action. Typically, the relevant facts about the various paths available to the agent would be accessible to him, and he would voluntarily choose a right action (rather than a morally objectionable one). Here we would say that the agent avoided blameworthiness; but this is a very different sort of context from the Frankfurt-type cases. In the Frankfurt-type cases, the agent does not choose to be morally responsible rather than not – these issues play no role in his deliberations. And in the alternative scenario in a Frankfurt-type case, the agent does not choose to escape responsibility or voluntarily choose anything which implies his escaping responsibility. (1999, pp. 120-1)

Of course, we are now taking it for granted that the agent does not know that they are in a Frankfurt-style case. But note that Fischer claims that where Jones avoids blame, he would be avoiding blame 'by accident' (1999). It is because Jones would be avoiding blame 'by accident' or by luck that Fischer thinks it cannot be said of Jones that he has the ability to avoid blame. For instance, in the counterfactual sequence Jones might have been about to decide to shoot the first

lady rather than the president. If this were the case then by hypothesis intervention would occur, ensuring that Jones decides to shoot the president. Jones would thereby have avoided blame, but ‘by accident’. He was not intending to avoid blame, but did so nevertheless. Jones is blameless, but by luck, and it is therefore not to Jones’ credit that his moral slate is unsullied by this episode.

But if this is the point that Fischer is making, then his argument is fallacious if it is supposed to show us why PAB is refuted. For it does not follow from the fact that one achieved something by accident, that one did not have the power to achieve it. Oliver, for example, has a device that enables him to ensure that a certain slot machine pays out. On this particular occasion however, Oliver is playing the slot machine, but has no interest in getting it to pay out and so does not use his device. The possibility of the machine paying out plays no role in his continuing to play the machine - he’s just idling away time, and is not playing to win. As it happens though, the machine does pay out. In this case it was a matter of luck that the machine paid out, but it is also true that Oliver had the power to make it do so. The fact then, that something happens by accident or by luck, does not entail that the agent did not have the power to bring it about that it happens. Return now to Jones. If Jones had been bent on avoiding blame, then he would have been bent on making some decision other than the decision to shoot the president. In this case, by hypothesis Jones would have been made to decide to shoot the president, and would thereby have avoided blame. In this case Jones exercises his power to avoid blame – he wanted to avoid blame and did avoid blame. But in those cases in which Jones is not bent upon avoiding blame – where avoiding blame plays no role in his deliberations, then Jones still might avoid blame – it still might be the case that intervention takes place, because, say, Jones was about to decide to shoot the first lady and not the president. But this does not show that Jones lacked the ability to avoid blame. I can simply agree with Fischer then, that the agent might avoid blame without it having played any role in their deliberations whatsoever – the agent might avoid blame by luck or accident. But this is consistent with it being the case that they had the ability to avoid blame.

The point can be made in a slightly different way if we return to Frankfurt-style cases in which Jones is fully aware of the situation that he is in. We have already agreed, I hope, that in such a case Jones does have the ability to avoid blame. Yet it is still a possibility that Jones will avoid blame by accident. He might, for instance, die suddenly of a stroke or alternatively the counterfactual intervener might accidentally intervene making Jones decide to do what in fact he was just about to freely decide to do. Such possibilities should not lead us to conclude that Jones does not, after all, have the ability to avoid blame. To reiterate, the fact that something happened by accident does not entail that one lacked the ability to bring it about. The fact that something was avoided by accident, does not entail that one lacked the ability to avoid it.

But perhaps Fischer’s point is a different one. Perhaps he is claiming that Jones lacks

the ability to avoid blame because he cannot do anything to ensure that he freely acts blamelessly, rather than being blameless in virtue of the occurrence of something not under his control. But how does this show that Jones lacks the ability to avoid blame? According to my definition of PAB, it is a condition of blame that the agent have the ability to avoid blame – nothing is said about an ability to incur blame. Perhaps Fischer is assuming that one does not have the ability to ensure that something does not happen, unless one has the ability to ensure that it does happen. If so, the claim is false. Consider that an agent can have the ability to prevent some event occurring, even where they lack the ability to ensure that the said event occurs. I may have in my possession a electronic device which enables me to ensure that any incoming missiles about to hit my home are destroyed. I therefore have the ability to ensure that my home is not destroyed by missiles. But it does not follow that I have the ability to ensure that my home is destroyed by missiles, for it is outside my control whether any missiles are directed at my home. Understood as such Fischer's criticisms are really criticisms of a different principle - the principle of avoidable and ensuring blame. This principle (and I know of no-one who explicitly defends it) states that it is a condition of being blameworthy that the agent have both the ability to avoid incurring blame and the ability to incur blame. Now I accept that such a principle is refuted by Frankfurt-style cases, but I do not defend such a principle. What is important (and the reason, I claim, that we care about PAP) is that we have the ability to keep our slates clean, not that we have the ability to ensure that they acquire dirt. None of Fischer's criticisms above suggest that we lack that ability in a Frankfurt-style case.

So far I have argued that Frankfurt-style cases in which the agent is fully aware of their situation refute PAP. Significantly though, such cases also show that the ability to avoid blame does not presuppose the ability to do otherwise. PAB does not entail PAP. This by itself would not show that PAB is valid, for PAB might nevertheless be refuted by Frankfurt-style cases in which the agent is unaware of the Frankfurt-style nature of their situation. However I went on to argue that Frankfurt-style cases do not refute PAB, for even though there is the possibility of the agent avoiding blame by accident or by luck, this in itself does not imply that they lacked the ability to avoid blame. Therefore, whilst Frankfurt-style cases refute PAP, they do not refute PAB.

What I now want to argue is that PAB captures all that is important about PAP. In other words, we should be unconcerned that PAP is refuted, so long as can keep PAB.

Consider then, that a Frankfurt-style case refutes PAP because it presents us with a case in which it is intuitive to consider the agent blameworthy for some decision they have made, even though they could not have decided otherwise. But I hold that a Frankfurt-style case in which the agent could have decided otherwise, might be equally effective at refuting PAP. Imagine that the

counterfactual intervener will make Jones decide to assassinate the president unless it is evident to him that Jones is about to either decide to assassinate the president on his own, or is about to decide to shoot the first lady on his own. In this Frankfurt-style case it is inevitable that Jones will make a morally objectionable decision, but if Jones freely decides, say, to assassinate the president in the actual sequence, it is true that Jones could have decided otherwise – he could have freely decided to assassinate the first lady. I claim that this sort of Frankfurt-style case is just as effective at refuting PAP. For it is implausible that a PAP defender will be satisfied with just any alternative possibility. A PAP defender wants alternative possibilities which include some for which the agent would not incur blame for having chosen. Note, one might object at this point that I am assuming that a PAP defender will necessarily reject the possibility of moral dilemmas. In other words, that I am assuming that a PAP defender rejects the possibility of situations arising in which an agent has to make one of two choices, both of which are morally impermissible. But this would be mistaken. For moral dilemmas concern moral wrongdoing, not moral responsibility.

If this is right, then the only reason we thought moral responsibility required alternate possibilities was because we thought it was required if the agent was to have the ability to avoid incurring blame. But if what I have been arguing has been correct, then we need not worry, for in fact alternate possibilities are not a necessary requirement of an ability to avoid blame. Frankfurt-style cases actually help show that an agent can have the ability to avoid blame, even if they lack completely the ability to do anything other than what they do in the actual sequence.

Notes

¹ Frankfurt's counterexample is actually a development of a case presented by John Locke: a man be carried whilst fast asleep into a room where is a person he longs to see and speak with, and be there locked fast in, beyond his power to get out; he awakes and is glad to find himself in so desirable company, which he stays willingly in. (1690 bk 2 chap 11 sec 10)

² Recent criticisms have been made by Kane, Widerker, Wyma and Ginet, all of whom argue that Frankfurt-style cases are question begging against the incompatibilist about determinism and moral responsibility (see Kane 1996; Widerker 1995; Ginet 1996; Wyma 1997). The criticism has not gone unanswered (see Haji 1998; Mele and Robb 1998; Stump 1999; Fischer 1999 and Hunt 2000). Another approach has been to highlight certain costs of rejecting PAP. Widerker and Copp have both argued that to reject PAP is to be committed to rejecting the Kantian principle that ought implies can (Widerker 1991; Copp 1997). But again, there have been responses both from those who deny that this is the case, such as Zimmerman, Haji and Yaffe, and from those who think it is true, but think that this is so much the worse for the Kantian principle (Fischer 1999 and 2003).

References

- Aristotle. 1915: *Nichomachean Ethics*. Vol 9 of *The works of Aristotle*. Ed. By W. D. Ross. London: Oxford University Press
- Copp, D. 1997. "Defending the Principle of Alternate Possibilities: Blameworthiness and Moral Responsibility" *Nous* 31 pp. 441-56
- Kane, R. 1996. *The Significance of Free Will* New York: Oxford University Press
- Fischer, J. M. 1994: *The Metaphysics of Free Will* Blackwell: Cambridge
- Fischer, J. M. 1999: Recent Work on Moral Responsibility in *Ethics* 110 pp. 93 – 139
- Fischer, J.M. 2003. "Ought Implies Can", causal determinism and Moral Responsibility" *Analysis* 63 pp. 244-50
- Frankfurt, H. 1969. "Alternate Possibilities and Moral Responsibility" *Journal of Philosophy* 66 pp. 829-39
- Ginet, C. 1996. "In Defence of the Principle of Alternative Possibilities: Why I Don't Find Frankfurt's Argument Convincing" *Philosophical Perspectives* 10: 403-17
- Haji, I. 1998: *Moral Appraisability*. New York: Oxford University Press.
- Hunt, D. 2000: *Moral Responsibility and Unavoidable Action* *Philosophical Studies* 97.2 pp. 195-227
- Kant, I. 1960. *Religion within the Bounds of Reason Alone*. Trans. By T. Greene and H. Hudson. New York: Harper and Row
- Locke, J. [1690] 1965: "An Essay Concerning Human Understanding. Abridged and ed. Maurice Cranston. New York: Collier
- McKenna, M. 1997. *Alternative Possibilities and the Failure of the Counterexample Strategy*. *Journal of Social Philosophy* 28: 71-85
- Mele and Robb. 1998. "Rescuing Frankfurt-Style Cases". *The Philosophical Review* 107.1 pp. 97-112
- Otsuka, M. 1998. "Incompatibilism and the Avoidability of Blame" *Ethics* 108 pp. 685-701
- Sinnott-Armstrong, W. 1988: *Moral Dilemmas*. Oxford: Blackwell Publishers.
- Stump, E. 1999. *Dust, Determinism, and Frankfurt: A Reply to Goetz*. *Faith and Philosophy* Vol 16 no 3 pp. 413-422
- Widerker, D. 1991: "Frankfurt on 'Ought Implies Can' and Alternative Possibilities" *Analysis* 51: pp. 222-24
- Widerker, D. 1995. "Libertarianism and Frankfurt's Attack on the Principle of Alternative Possibilities" *Philosophical Review* 104 pp. 247-61
- Wolf, S. 1980: "Asymmetrical Freedom" *The Journal of Philosophy* 77 pp. 151-166
- Wyma, K. 1997. "Moral Responsibility and Leeway for Action" *American Philosophical Quarterly* 34 pp. 57-70
- Yaffe, G. 1999: "Ought Implies Can' and the Principle of Alternate Possibilities" *Analysis* 59 pp. 218-22

Rec.: 17/03/2004

Ap.: 08/07/2004

e-mail: g.k.harrison@durham.ac.uk