

Frankfurt-Style Cases and Unavoidable Blameworthiness

Shohei TAKASAKI

Abstract

Frankfurt's well-known counterexample to the Principle of Alternative Possibilities has dramatically changed the landscape of debates about free will and moral responsibility. Though the discussion of Frankfurt-Style Cases is vast, one important question has not been adequately addressed: the question of whether someone can be unavoidably blameworthy. I will argue that examination of this issue raises an important challenge to the proponents of Frankfurt-Style Cases.

In Section 1, I argue that those who endorse Frankfurt-Style Cases must admit the possibility of justifiable unavoidable blame. In Section 2, I examine Widerker's objection to the possibility of unavoidable blameworthiness. In Section 3, I examine McKenna's reply to Widerker and his defense of unavoidable blameworthiness. In Section 4, I argue that the disagreement between Widerker and McKenna is due to the fact that they focus on very different aspects of moral responsibility. I use Gary Watson's well-known distinction between the "accountability" and "attributability" senses of responsibility to propose that Widerker and McKenna are talking at cross purposes. Widerker should be understood as arguing that accountability requires avoidability, whereas McKenna is claiming that attributability does not require avoidability. I will argue that an agent is morally blameworthy only if she is accountable.

1. Frankfurt-Style Cases as Cases of Unavoidable Blame

In the literature before Frankfurt's influential paper (Frankfurt 1969), one principle of moral responsibility, the Principle of Alternative Possibilities (henceforth, **PAP**), had been largely assumed. This principle can be stated roughly as follows:

PAP: The agent S is morally blameworthy for his action φ only if S could have done otherwise than φ ¹.

Intuitively, **PAP** seems highly plausible. Suppose, for example, that while you drive a car you bring about a traffic accident resulting in injury. Your friend would blame you for what you did. But what if a pedestrian abruptly ran out in front of your car from a place where normally there would not have been anyone, and there was no way you could have avoided hitting him? In that case, you might be regarded as unlucky, and your friend would withdraw his original judgment that you are blameworthy. It seems reasonable to think that the difference between this case and a traffic accident from one's carelessness lies in whether or not the agent could have done otherwise, as **PAP** states. Moreover, we often excuse ourselves by saying that we could not have done otherwise.

Let us move on to the basic idea of Frankfurt-Style Cases: that they are counterexamples to **PAP**. Here is one of the instances:

Jones is deliberating whether or not to keep a certain promise he made to his uncle. Unbeknownst to Jones, there is another person, Black, who for some reason does not want Jones to keep his promise. Black has the power and the means to force Jones to break the promise. But wishing to avoid showing his hand unnecessarily, he has made up his mind to intervene if and only if Jones does not show a sign of going to decide to break the promise. Call that sign 'S1'. If Jones does show that sign, then Black does nothing, knowing that in this case Jones will act as he (Black) wants him to act. (It is assumed that Black knows Jones very well in this regard) Finally, suppose that Jones decides to break the promise for reasons of his own. (Widerker 2003, p. 53)

In this case, Jones seems to have no alternative possibilities. Even if Jones tried to keep his promise, Black would intervene and make sure that Jones would do what Black wants him to do. Thus, Jones could not have avoided deciding to break his promise. Or so it is argued.

Then, if Jones *is* morally blameworthy for his action, the case above refutes **PAP**.

All Frankfurt-Style Cases depend on the following two assumptions, as far as they are designed to show that **PAP** is false:

IRR: There may be circumstances that in no way bring it about that a person performs a certain action; nevertheless those very circumstances leave him with no morally significant alternative to performing that action. Let us call this kind of case an IRR-situation².

MB: In an IRR-situation, the agent is morally blameworthy for the morally wrong act he performed.

I define *FC-Defenders* as those who accept both **IRR** and **MB**, and I will argue against FC-Defenders by examining the assumption **MB** carefully. Though whether or not **IRR** is true is a quite controversial metaphysical issue which is worth pursuing, I concede for the sake of argument that an IRR-situation is metaphysically possible. Hereafter I pretend to take Jones' case above as an IRR-situation.

My objection to FC-Defenders has two steps:

- (i) **IRR** and **MB** entail that there are cases where blame is unavoidable yet justifiable.
- (ii) There are no good reasons to suppose that there are cases where blame is unavoidable yet justifiable.
- (iii) Therefore, given the assumption of **IRR**, **MB** is false.

I will support (ii) in the subsequent sections. In the rest of this section, I give a short argument for (i). Suppose that Jones' case is an IRR-situation. Then, Jones could not have avoided breaking his promise. And, it follows from this together with **MB** that Jones could not have acted in a way such that he is not blameworthy. Therefore, Jones could not have avoided being blameworthy for his action. This is an instance of unavoidable blameworthiness.

2. Widerker's W-Defense

Suppose, for the sake of argument, that **IRR** is true. And suppose also that in some particular situation which satisfies **IRR**, Jones decided to break his promise *on his own* with no intervention of the counterfactual intervener Black. Is Jones morally blameworthy for his action in this situation? As I showed in the previous section, FC-Defenders answer “Yes” to this question, and this entails that they accept the claim that unavoidable blame is justifiable. But this claim is far from obvious and some philosophers have argued against it³. I will focus on Widerker's “W-defense” argument.

Widerker (2003) gave a compelling reason to think that Jones is *not* morally blameworthy for his action. He begins his argument by asking FC-Defenders the following question:

Let me grant, for the sake of discussion, that in the IRR-situation under consideration, Jones acted freely in the sense that what he did he did for reasons of his own without being causally determined or coerced to so act. Still, since you, Frankfurt, wish to hold him blameworthy for his decision to break his promise, tell me *what, in your opinion, should he have done instead?* Now, you cannot claim that he should not have decided to break the promise, since this was something that was not in Jones's power to do. Hence, I do not see how you can hold Jones blameworthy for his decision to break the promise. (Widerker 2003, p.63)

This defense is called the “What-should-he-have-done defense” or “W-defense” for short. In the passage quoted above, Widerker is claiming the following two propositions:

- (i) There is no proper answer to the “what should he have done?” question.
- (ii) If there is no proper answer to that question, then holding Jones blameworthy is illegitimate.

Let us take a look at (i) first. Though it is not entirely clear how Widerker defends this claim, I think he has in mind the following kind of reasoning. Suppose that some FC-Defender claims that Jones should have done A, where A is a different action than the actual one. Then, either Jones could have done A or could not have done A. Suppose, first, that Jones could have done A. Then, contrary to the assumption that Jones' case is an IRR-situation, Jones seems to have genuine, morally relevant alternatives; they are obviously morally relevant because one is supposed to be a blameworthy action and the other is not. But since FC-Defenders endorse **IRR**, they cannot take this option.

Second, suppose that Jones could not have done A. How does Widerker deduce (i) from this supposition⁴? His reasoning can be formulated as follows:

1. Jones could not have done A.
2. If an agent *S* could not have done an action φ , then it is morally unreasonable to expect *S* to perform φ ⁵.
3. It is morally unreasonable to expect Jones to perform A. (by (1) and (2))
4. To hold that *S* should have done φ is to expect *S* to do φ .
5. It is morally unreasonable to hold that Jones should have done A. (by (3) and (4))
6. Therefore, the answer “he should have done A” is not a proper answer to the question.

Since this argument holds for any arbitrary action A, we can conclude that there is *no* proper answer to the “what should have done” question.

Substantial premises of the argument above are (2) and (4). Though I cannot defend these premises fully here, let me note some points. First, premise (2) looks highly plausible, since it seems morally unreasonable to expect an agent to perform an action that is impossible for him. Second, (4) reflects what Widerker intends to mean by asking “what *should* he have done?”. He regards “should have” statements as expressing the demand for an agent to do something. Some FC-Defenders might resist Widerker's interpretation and claim instead

that “should have” statements express a moral obligation an agent has in the situation.

Let us move on to (ii). Widerker defends (ii) by appealing to the following constraint on ascriptions of moral blame:

PAE: An agent *S* is morally blameworthy for doing φ only if under the circumstances it would be morally reasonable to expect *S* not to have done φ . (ibid, p. 63, notations slightly altered)

Since our consideration of (i) tells us that it is morally unreasonable to expect Jones not to have broken his promise, by **PAE** we can conclude that Jones is not morally blameworthy for breaking his promise.

PAE is controversial and not so clear on its own, though. For one thing, as Zimmerman argues, Widerker has not made explicit the exact meaning of “be morally reasonable to expect”. That said, I find Widerker’s understanding of moral blame compelling. I think that **PAE** might be supported further if we have a look at an essential feature of the practice of holding blame. According to G. Watson (1996), “[j]udgments of moral blameworthiness [...] involve the idea that agents deserve adverse treatment or “negative attitudes” in response in their faulty conduct” (Watson 1996, p. 266). That is to say, a judgment about an agent’s blameworthiness essentially involves the thought that the agent deserves to suffer some harm or sanction from other people. Then, if **PAE** is false, there seem to be cases where the agent is inevitably liable to harm or sanction from other people. But how can such a thing be justified? I can think of no good answer to this question. In the absence of some compelling argument against **PAE**, I am inclined to accept it.

3. McKenna’s L-Reply

In the last section, we examined one argument against unavoidable blameworthiness from Widerker and showed that the conception of moral blame Widerker relies on has considerable plausibility. In this section, we will consider McKenna’s reply (“L-reply”) to

Widerker's W-defense. I think contrasting Widerker's W-defense and McKenna's L-reply is very important, because their disagreement suggests that they have a substantial disagreement about how to understand the notion of moral responsibility.

McKenna concedes the dialectical power of the W-defense and admits that there is no proper answer to the "what should he have done?" question. Instead of giving a direct answer, he replies in an indirect way: "Look at what the agent has done." (McKenna 2008, p.785) McKenna argues that by looking at the actual sequence of the agent's action, we can legitimately hold that the agent is blameworthy. McKenna's L-reply is important, for it effectively expresses the reason FC-Defenders think that Jones in an IRR-situation is morally blameworthy for his action.

It is worth examining carefully how McKenna and other FC-Defenders hold Jones blameworthy in our scenario. What respect of Jones's actual sequence tells us that he is blameworthy? Frankfurt (2003), for instance, gives a concise answer: "We claim [...] that it may be entirely reasonable to blame a person for having done something that he cannot reasonably be expected to have avoided doing. In our judgment, this may be reasonable when the person has performed the action in question — like Jones — entirely for reasons of his own and thus of his own free will" (Frankfurt 2003, p. 344). I think most FC-Defenders share Frankfurt's idea that Jones is blameworthy because he decided to break his promise *on his own*. Then we can further ask two questions to FC-Defenders: First, when is one's action regarded as one's own? Second, how is that "ownership" condition connected to the notion of moral blameworthiness? I will examine how FC-Defenders would respond to each question in turn.

Consider first the question about the ownership condition. To my surprise, many FC-Defenders seems satisfied to claim only that Jones is blameworthy because he performed his action on his own, leaving the explanation of what "on one's own" means unstated. I suppose one natural answer would be that one's action is one's own only if that action (or the reasons on which that action is based) expresses what the agent stands for or expresses the agent's moral character. This proposal assumes that each action, or more importantly, each decision and choice, expresses how the agent is. This assumption seems highly

plausible when we reflect on our everyday decisions or choices. Of course, an action is not necessarily a proper expression of the agent's character. But in such cases, the agent would have some *excuse* for his action and that excuse would tell us how the agent really is. Moreover, the proposal above accommodates the fact that in some irregular cases, such as manipulation, coercion, or addiction, even if the agent's action is morally vicious, we would not hold the agent blameworthy. For it might be said that in these cases the action is not the agent's own in the morally relevant sense⁶.

Let us move on to the second question, that is, the question about the relation between the ownership condition and moral blameworthiness (in a case of culpable action). Since we have seen that FC-Defenders focus on what the action discloses about the agent's moral character, it seems that they believe that to hold someone blameworthy is to make a negative evaluation of that person's moral character. Then the necessary condition for moral blameworthiness can be stated roughly as follows: The agent *S* is morally blameworthy for his action ϕ only if the action reveals something morally bad about *S*'s moral character.

Given the considerations so far, we can state the FC-Defenders' understanding of moral blameworthiness as follows:

FCD-BL: The agent *S* is morally blameworthy for his action ϕ if and only if

- (i) ϕ or the reasons on which ϕ is based expresses *S*'s moral character, and
- (ii) ϕ reveals that *S*'s moral character is in some way bad.

My argument has not shown that there are no other viable options available for FC-Defenders than **FCD-BL**. Nevertheless, I do not see how such other conceptions of moral blameworthiness are open to them, given that FC-Defenders focus exclusively on the actual sequence of the agent's action in order to evaluate his moral blameworthiness.

Given **FCD-BL**, it is easy to see how unavoidable blameworthiness is justified in the framework of FC-Defenders. According to them, an agent may be morally blameworthy for his action regardless of whether the action is unavoidable for him, provided only that the action expresses his moral character and the moral quality of this character is bad. I will

argue in the next section that this view is not completely friendly with our moral intuitions.

4. Attributability and Accountability

We have seen that Widerker and McKenna (and other FC-Defenders) hold very different conception of moral responsibility. How should we understand this disagreement? I do not want to say that either view is wrong; rather, I think that each view correctly focuses on a different aspect of moral responsibility.

I suggest that these aspects correspond to Gary Watson's well-known distinction between accountability and attributability, both of which, Watson argues, have a great significance in our moral practice. According to Watson, moral responsibility in the accountability sense is directly related to our practice of holding one another accountable for our actions, thereby requiring or demanding that their actions conform to certain standards. behavior of an agent. On the other hand, attributability concerns *aretaic* aspects of moral responsibility. Watson says that the aretaic perspective is concerned "with the question of what activities and ways of life are most choiceworthy" (Watson 1996, p. 266). We can make these judgments without thinking that we have the right or authority of "calling anyone out" if they fail to live up to the standards that we think choiceworthy.

Watson gives an example that nicely illustrates the distinction between accountability and attributability⁷. Suppose some philosopher has written a paper in which he denies the reality of moral responsibility. And yet this same philosopher, when some of his books were stolen, said, "That was a shoddy thing to do". For those who fail to recognize the distinction between accountability and attributability, what he said might seem inconsistent. But once we acknowledge that distinction, we can see that his claim is not inconsistent: to call the conduct "shoddy" is to attribute some negative moral evaluation to the person on the basis of the conduct; to deny the reality of moral responsibility is to deny the reality of accountability, that is, to deny that anyone can ever be fairly and legitimately *blamed* for any of his actions.

Now, in view of Watson's distinction between the two faces of moral responsibility, we

can make sense of the disagreement between Widerker and McKenna. On the one hand, Widerker clearly has in mind an accountability sense of moral responsibility, because he directly addresses the question of whether it is fair to hold Jones blameworthy. On the other hand, McKenna and other FC-Defenders are thinking, whether they realize this or not, of an attributability sense of moral responsibility — because their arguments are based on the aretaic evaluation of the quality of the agent's moral character.

If this diagnosis is cogent, FC-Defenders face a significant challenge to their views when it comes to the question of moral blameworthiness. The problem is that, though FC-Defenders would willingly hold Jones morally responsible, it is far from obvious that their conception of moral responsibility is relevant to the notion of blameworthiness. For they justify their claim by appealing to the attributability sense of responsibility, not the accountability sense. To put it another way, there seems to be a good reason to doubt the truth of **FCD-BL**, insofar as it is construed as the condition of blameworthiness.

There are several reasons to doubt that attributability is relevant to the agent's blameworthiness. First, as Zimmerman (2003) points out, judgments concerning attributability (which he calls *aretaic* judgments) and judgments concerning accountability (which he calls *hypological* judgments) arguably come in different categories among the broader notion of moral judgment, and they are at least logically independent; that is, judgments of attributability (say, that he is shoddy) do not entail those of accountability (say, that he is blameworthy), and vice versa. Thus FC-Defenders have to offer some “bridge” principle between these two notions.

Second, I think the irrelevance of attributability to blameworthiness becomes more vivid when we consider the cases of inevitable evil character, or more generally, the problem of “constitutive luck” (Nagel 1976). Suppose some agent committed a murder on his own. He retained all the capacities FC-Defenders would require for moral responsibility. But suppose also that it turned out that his morally vicious character had been formed by some genetic factors and the environment when he was young (say, he suffered from terrible physical abuse from his parents), which are totally beyond his control. He, after all, could not avoid having his own evil character. Then, how do we respond to his conduct after we know his

terrible young period? It seems quite natural to have an “ambivalent” response; some of us would be reluctant to hold him blameworthy. It is not so obvious how FC-Defenders can properly explain this mixed feeling⁸.

FC-Defenders might reply that the murderer is still blameworthy, even though his evil character is unavoidable, insofar as he himself wholeheartedly *identifies* with or *endorses* his character (cf. Frankfurt 1975). But I don't find this reply convincing. The agent's identification with his own character might entail that he willingly accepts the corresponding negative attitudes or sanctions he would suffer from others, but that does not entail that holding him blameworthy is legitimate. The question of whether it is fair or legitimate to blame someone should not be a matter of subject's personal feeling, but rather, matter to be considered from more objective point of view.

5. Conclusion

I have shown that the disagreement between Widerker and McKenna is due to their different conceptions of moral responsibility: accountability and attributability. If this diagnosis is correct, then a proper answer to the question “Is Jones morally responsible for his action?” would be “Yes and no”. But when it comes to the question of his moral blameworthiness, the proper answer would be “No”, on the assumption that **IRR** is true in the Jones's scenario. This consequence, if correct, would be significant because, so far as I can see, no FC-Defenders have given us a good reason to think that Jones is *blameworthy* for his action. To do this, they would either have to show that they are not talking about the attributability sense of moral responsibility, or offer some bridge principle between attributability and blameworthiness⁹.

¹ Here I limit my concern to moral blameworthiness, rather than moral responsibility more generally. This is not uncontroversial, but I take it for granted since both proponents and opponents of Frankfurt-Style Cases will not deny that what they are talking about is moral blameworthiness.

² I borrowed this from Widerker (2006), though the original name of this statement is (G-IRR).

³ For example, see M. Otsuka (1998) and D. Copp (2003).

⁴ One natural way would be to appeal to “Ought Implies Can” Maxim. But as Capes (2010) points out, Widerker denies that his argument turns on the Maxim. For the debate over the relation between the Maxim and Frankfurt-Style cases, see D. Copp (2003) and G. Yaffe (1999).

⁵ Here *S* and *φ* are meant to be variables, whereas *A* describes some specific action.

⁶ FC-Defenders might further argue that, following Frankfurt’s own view on moral responsibility (Frankfurt 1975), in order for an action to be genuinely on one’s own, the agent himself must endorse or identify one’s character (or in Frankfurt’s term, “deep self”). In other words, it might be argued that a sort of agent’s active commitment to one’s own character is needed.

⁷ The original example comes from van Inwagen (1983).

⁸ Of course, FC-Defenders have made an enormous effort to address this issue, particularly through replying to the well-known “manipulation argument” (cf. Pereboom 2001). They would argue that the history of how murderer’s character was formed is so unusual (in the sense which could be comparable with covertly manipulated cases) that he is not blameworthy. This is a difficult matter I cannot address here.

⁹ This paper is sponsored by Japanese Society for the Promotion of Science (17J07640).

References

- Capes, J. 2010. “The W-Defense”, *Philosophical Studies* 150, 61–77.
- Copp, D. 2003. “‘Ought’ Implies ‘Can’, Blameworthiness, and the Principle of Alternate Possibilities” in D. Widerker & M. McKenna (Eds.), in *Moral Responsibility and Alternative Possibilities* (pp. 265–299). Aldershot, UK: Ashgate Press.
- Frankfurt, H. 1969. “Alternate Possibilities and Moral Responsibility”, *The Journal of Philosophy* 66, 829–839.
- ———. “Some Thoughts Concerning PAP”. in D. Widerker & M. McKenna (Eds.), in *Moral Responsibility and Alternative Possibilities* (pp. 339–348). Aldershot, UK: Ashgate Press.
- McKenna, M. 2008. “Frankfurt’s Argument against the Principle of Alternative Possibilities: Looking beyond the Examples”, *Nous* 42, 770–793.
- Nagel, T. 1976. “Moral Luck”, in *Proceedings of the Aristotelian Society* 50, 137–151.
- Otsuka, M. 1998. “Incompatibilism and the Avoidability of Blame”, *Ethics* 108, 685–701.
- Pereboom, D. 2001. *Living Without Free Will*. Cambridge: Cambridge University Press.
- van Inwagen, P. 1983. *An Essay on Free Will*. Oxford: Clarendon Press.
- Watson, G. 1996. “Two Faces of Responsibility”, *Philosophical Topics* 24, 227–248.
- Widerker, D. 2003. “Blameworthiness, and Frankfurt’s Argument against the Principle of Alternative Possibilities”, in *Moral Responsibility and Alternative Possibilities* (pp. 53–73). Aldershot, UK: Ashgate Press.
- ———. 2006. “Libertarianism and the Philosophical Significance of Frankfurt Scenarios”, *Journal of Philosophy* 103(4), 163–187.
- Yaffe, G. 1999. “‘Ought’ Implies ‘Can’ and the Principle of Alternate Possibilities”, *Analysis* 59, 218–222.
- Zimmerman, M. J. 2003. “The Moral Significance of Alternate Possibilities”, in *Moral Responsibility and Alternative Possibilities* (pp. 301–325). Aldershot, UK: Ashgate Press.