

The W-Defense

Penultimate Draft

Final version in *Philosophical Studies* 150 (2010): 61-77

The final publication is available at link.springer.com

By Justin Capes

Harry Frankfurt has famously argued that the Principle of Alternative Possibilities, or PAP, according to which a person is morally responsible for what she has done only if she could have done otherwise, is false.¹ According to Frankfurt, there may be circumstances which make it impossible for a person to avoid performing a certain action but that in no way bring it about that the person performs that action, and he contends that in circumstances of this sort the agent may be morally responsible for what she has done despite the fact that she couldn't have done otherwise. There has been a great deal of critical discussion of this argument in recent years, virtually all of which has focused on the question of whether the sorts of circumstances envisioned by Frankfurt (FR-situations) are possible. Until quite recently, however, very little attention has been given to the question of whether an agent in FR-situations could ever be morally responsible for what she has done. One critic who has addressed this question is David Widerker (2000, 2003, and 2005). He contends that even if FR-situations are possible, Frankfurt's argument does not succeed against the following version of PAP for blameworthiness:

PAPb: A person is morally blameworthy for what she has done only if she could have done otherwise.

(Henceforth when I refer to PAP I shall be referring to this version of the principle.) According to Widerker, agents in FR-situations could never be morally blameworthy for what they have done. For reasons that will become clear shortly, he refers to the argument for this claim as the what-should-he-have-done-defense, or the W-defense for short.

The W-defense is important for at least two reasons. First, as I explain in a bit more detail in section one, the dispute over the possibility of FR-situations has proved difficult to resolve, and if current trends are any indication, no resolution of the matter is forthcoming. However, the W-defense suggests a way of advancing the debate over PAP which would allow one to bypass the thorny metaphysical issues involved in assessing the possibility of FR-situations. Secondly, the W-defense raises interesting questions regarding the nature of moral blame, and as I argue towards the end of the essay, one's

¹ It is best to think of PAP as restricted to direct, as opposed to derivative, responsibility even though the principle is not explicitly set out as such. An agent might be indirectly (or derivatively) responsible for performing some action that she couldn't avoid performing, if her performance of that action is the result of previous acts of hers that were avoidable. So, to take the classic example, a drunk driver may not be able to avoid hitting a pedestrian given his intoxicated state. This does not falsify PAP, however, since in typical cases the driver is derivatively responsible for hitting the pedestrian. In most cases, the driver's responsibility for hitting the pedestrian derives from his responsibility for previous acts of his—for example, his decision to drink and drive. The question at issue with respect to PAP, however, is whether an agent might be directly (or non-derivatively) responsible if she is unable to do otherwise.

assessment of Frankfurt's argument against PAP will depend to some extent on what one thinks is involved in holding a person blameworthy for what she has done.

In what follows I examine the W-defense in some detail, along with three recent replies to it by defenders of Frankfurt's argument, including Frankfurt himself. I contend that each of these replies is problematic and, indeed, that two of them play directly into the hands of the PAP-defender. I then develop my own reply in defense of Frankfurt by calling into question an assumption which is at the heart of the W-defense regarding the nature of moral blame. I begin, however, by briefly locating the W-defense within the larger dialectic surrounding Frankfurt's argument.

1 Frankfurt-scenarios, dialectical stalemate and the W-defense

Frankfurt's argument against PAP relies on the following two assumptions:

- FR: There may be circumstances that make it impossible for a person to avoid performing a certain action but that in no way bring it about that the person performs that action.
- M: It is possible that a person who performs a morally wrong action in a FR-situation is morally blameworthy for what she has done.

It follows from these two claims that PAP is false. As I mentioned in the introduction, the controversy surrounding this argument has focused almost entirely on the truth of FR and whether the circumstances described therein are possible. This is where the infamous Frankfurt-style examples become relevant. To illustrate a FR-situation, Frankfurt offered this example:

Black ...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... [However] 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 (1969: 835-36).

To fill out the example, let us suppose that the action Black wants Jones to perform is shooting Smith, and let us refer to the example as "Shoot Smith." This is supposed to be a FR-situation: given the presence of Black, Jones is apparently unable to do otherwise than shoot Smith, but Black does not force, coerce or otherwise induce Jones to act. Jones shoots Smith entirely on his own, and for this reason many find the claim that Jones is morally blameworthy intuitively quite appealing.

Many PAP-defenders have challenged Frankfurt's argument by questioning whether this example really is a FR-scenario. If it is not, then Frankfurt will have failed to establish a key premise of his argument. And there are indeed a number of problems with

the example. For starters, we might wonder how Black could possibly determine what Jones is going to do before he actually does it. One way of answering this question is to imagine that Black is a particularly gifted neuroscientist who has developed sophisticated means of observing Jones's neural states from afar and that on the basis of his observations is able to detect a certain neural pattern or sign which he knows is somehow correlated with decisions to perform the sort of action he wants Jones to perform. Thus, if at any point during the deliberative process Jones exhibits this sign, Black will know that intervention on his part is unnecessary. This way of embellishing the case is somewhat fanciful to be sure, but all Frankfurt needs in order to make his point is a metaphysically possible scenario, and this version of the example, fanciful as it may be, might seem to turn the trick.

This response gets the example out of the frying pan only to put it into the fire, however, for as several PAP-defenders have pointed out, the story now faces the following dilemma: either the relation between the prior sign and Jones's subsequent action is deterministic or it is not. If it is, then claiming that Jones is blameworthy for what he has done would beg the question against incompatibilists, who think that determinism precludes moral responsibility. If, on the other hand, the relation between the prior sign and Jones's action is not deterministic—if, for example, the sign is merely a reliable, but ultimately fallible, indicator of what Jones is going to do—then the presence of the sign will not guarantee that Jones is going to shoot Smith. In other words, if the prior sign is not causally sufficient for the outcome Black desires, the only way he can guarantee the desired outcome is to intervene. But since he does not intervene, and assuming Jones's action was not causally determined, there is no longer any reason to suppose that Jones was unable to do otherwise. Either way, then, the example will have failed to establish the possibility of FR-situations.²

Defenders of Frankfurt's argument have tried to avoid this dilemma by developing more sophisticated "Frankfurt-style" examples that either do not involve a prior-sign (e.g. Mele and Robb 1998, Hunt 2000, Stump 2003) or in which the relation between the prior-sign and the agent's action is such that it allows Black (or whatever plays the role of Black) to rule out all morally significant alternatives without causally determining the agent's action (Pereboom 2000, Hunt 2005).³ Whether any of the new Frankfurt-style examples succeed in avoiding the dilemma is still a matter of considerable controversy and no consensus on the matter has yet emerged.⁴

Suppose no consensus on the possibility of FR-situations can be reached. Would the debate over PAP then end in a stalemate? I should think not. Although the question of whether Frankfurt-style examples succeed in establishing FR is an important one, the

² Kane 1985, Widerker 1995, and Ginet 1996 have each developed versions of this dilemma.

³ It is generally agreed by all participants in the debate that no non-question begging example can be constructed that does not involve at least some alternative possibilities. The question, then, becomes whether the alternatives that remain in the new Frankfurt-style examples are morally significant, where a morally significant alternative (sometimes also called a 'robust alternative') is the sort of alternative that a defender of PAP would claim transforms a case from one in which the agent is not responsible to one in which the agent is responsible, and does so in such a way that the addition of the alternative would aid in explaining or "grounding" the claim that the agent is morally responsible. For my own part, I tend to think the remaining alternatives in the new examples are morally significant, but I cannot argue for this here.

⁴ My own view is that some of the new examples probably do succeed in avoiding the dilemma, though I shall not argue for this here.

intense focus on these examples has tended to obscure larger dialectical currents in the debate over PAP. In particular, it has eclipsed the question of whether agents in FR-situations could be morally blameworthy for their wrong actions as M claims. Many participants on both sides of the debate seem to have simply taken it for granted that M is true, but as we shall see, the matter is not beyond dispute. And if it could be shown that M is false, the PAP-defender would then be able to show that Frankfurt's argument fails quite independently of whether FR-situations are possible. This is significant, especially in light of the fact that the debate over the possibility of FR-situations seems to be moving towards a stalemate, if it has not reached that point already.

So let us turn our attention to M. In order to assess M there are several issues that would need to be taken into consideration which I shall not presently address in any detail. For instance, contrary to what much discussion of his argument might suggest, Frankfurt did not simply rely on the intuitive force of his examples to support the claim that agent's like Jones are responsible. He also offered two arguments in support of this claim. These arguments turn on the following two principles, respectively:

Irrelevance Principle. Facts that are irrelevant to explaining why an agent performs a certain action have no bearing on whether the person is morally responsible for performing that action.

Good Excuse Principle. A person can cite something as an excuse only if it reveals either that the person did not act from a culpable motive, or that the person's action issued in a causally deviant way.

Because his inability to do otherwise is irrelevant to an account of why Jones killed Smith—he would have killed him even if Black had not been present, that is, even if he could have done otherwise—it follows from the irrelevance principle that his inability to do otherwise has no bearing on his moral responsibility for shooting Smith. Moreover, the fact that he couldn't have done otherwise does nothing to show that Jones did not act from a culpable motive, nor does it show that his action issued in a causally deviant fashion (the action was apparently brought about in just the way it would have been brought about had Jones been able to do otherwise); hence, according to the good excuse principle, Jones cannot cite his inability to do otherwise as an excuse for shooting Smith. So if he is not morally responsible for shooting Smith, it cannot be because he couldn't have done otherwise but because of some other feature of the situation. Hence, whether he is responsible or not has nothing to do with his inability to do otherwise. These arguments are critically evaluated by Widerker (2000, 2003) and have recently been defended by Michael McKenna (2008). The issues raised here are interesting in their own right, but for now I shall put them aside. We should keep in mind, however, that a thorough assessment of M would need to take into account whether these arguments are defensible.

In the remainder of this essay I shall focus on the W-defense, which purports to give reasons why agents in FR-situations could never be morally blameworthy for what they have done; that is, it purports to show that M is false. The W-defense is perhaps one of the most significant responses to Frankfurt in recent years, not only because it suggests a way to advance the debate over PAP independently of the controversy surrounding FR,

but as I suggested in the introduction, it also raises interesting questions pertaining to the nature of moral blame. Indeed, it is precisely the assumptions about the nature of blame to which the W-defender is seemingly committed that make the argument ultimately untenable—or at any rate, so I shall be arguing.

2 Blameworthiness and the W-Defense

For the sake of argument, let us grant that Shoot Smith is a FR-situation: Jones is unable to avoid shooting Smith, and the circumstances that prevent him from doing otherwise do not bring it about that he shoots Smith. Having granted this, we must now ask whether Jones is morally blameworthy for what he has done. Widerker claims that he is not. In support of this claim he poses the following challenge to Frankfurt and other would-be defenders of Frankfurt's argument:

...since you...wish to hold Jones blameworthy for [shooting Smith], tell me *what, in your opinion, should he have done instead?* Now you cannot claim that he should not have [shot Smith], since this was something that was not in his power to do. Hence, I do not see how you can hold him blameworthy for [shooting him] (Widerker, 2003: 63).

Widerker is confident that no reasonable answer can be given to the question “What should Jones have done instead?” in which case he thinks we cannot legitimately hold Jones blameworthy, for as he goes on to explain

When we consider someone morally blameworthy for a certain act, we do so because we believe that morally speaking he should *not* have done what he did. This belief is essential to our moral disapproval of his behavior. Sometimes, however, such a belief may be unreasonable, for example, in a situation in which it is clear that the agent could not have avoided acting as he did (2003: 63).

According to Widerker, then, to hold Jones blameworthy would require that we believe he should have done otherwise, and to believe that a person should have done otherwise is unreasonable if the person is not able to do otherwise. Therefore, we cannot reasonably hold Jones blameworthy for what he has done even if, as Frankfurt claims, he acted of his own free will.

It is not immediately clear why Widerker deems it unreasonable to believe that an agent should not have performed some action if, given the circumstances, the agent could not have avoided performing that action. One might appeal here to the maxim that ‘ought’ implies ‘can’ (henceforth, the Maxim). In cases in which an agent could not have done otherwise, the Maxim implies that the agent was not obligated to do otherwise, and if an agent was not obligated to do otherwise, it is hard to see how one could reasonably claim that the agent *should* have done otherwise. However, Widerker cannot appeal to the Maxim here, for (in response to an objection raised by John Martin Fischer which I discuss shortly) he denies that the W-defense turns on the Maxim (2005: 303-304). What grounds Widerker's claim, rather, seems to be this: to believe that someone should not do something involves an expectation or demand that the person not do that thing. But in a situation in which it is clear to us that the agent could not have avoided performing some

action, it would be unreasonable to expect the person not to perform that action, since to do so would be to expect the person to do what for her, in those circumstances, is impossible. Hence, it would be unreasonable to believe that a person should not have done something she couldn't have avoided doing.

According to Widerker, the W-defense thus suggests the following constraint on moral blameworthiness, which he calls the Principle of Alternative Expectations:

PAE: An agent *S* is morally blameworthy for performing an action *A* only if under the circumstances it would be morally reasonable to expect *S* not to *A*.

Since it would clearly be morally unreasonable to expect an agent not to do something she is unable to avoid doing, PAE has the consequence that *S* is morally blameworthy for *A*-ing only if *S* could have avoided *A*-ing, and this principle just is (is identical to) PAPb. Consequently, the W-defense yields a comprehensive defense of PAP for moral blameworthiness.

To avoid confusion, we should distinguish between two different senses of the term 'expect'. Consider a statement like "I expect the storm to pass." In statements of this sort, the term is being used in an altogether different sense than it is being used in the W-defense or in PAE. To expect that the storm will pass is to anticipate or predict that the storm will pass. To expect that a person do something in the sense at issue in the W-defense, however, is to demand or insist that the person do that thing, as, for example, when one expects one's children to behave in public or when one expects one's students not to cheat on their assignments.

So Widerker's claim, if I have understood him correctly, is that while the agent in FR-situations like Shoot Smith has perhaps violated the demands of morality, it is unreasonable to hold that person blameworthy for his wrongdoing, for in doing so we would be placing an unreasonable expectation or demand on him. By holding him responsible we would be expressing an expectation or demand that he not have done something he was unable to avoid doing, and such an expectation clearly seems unreasonable.

Before turning to some recent responses to the W-defense, I should point out that there is an initial problem with the argument as Widerker formulates it. Contrary to what he claims, the W-defense does not directly support either PAE or PAP. To see this, notice that the conclusion of the W-defense is that we cannot legitimately *hold* an agent in FR-situations blameworthy. That it would be unreasonable to hold a person blameworthy, however, needn't imply that the person is not in fact blameworthy, for a person might be blameworthy for what she has done even if it would be unreasonable for others to hold her blameworthy. There might be cases in which a person is blameworthy but in which the available evidence suggests otherwise. In such cases it would be unreasonable for others to hold the person blameworthy and yet the person would be blameworthy nonetheless. *Holding* a person responsible, to the put the point in a slightly different way, is not the same as the person actually *being* responsible. A person might be held responsible for what she has done even though she is not in fact responsible, and she might be responsible even if no one actually holds her responsible or is justified in

holding her responsible. Even if it were successful, then, the W-defense would not directly suggest PAE but rather

PAE*: We can legitimately hold *S* morally blameworthy for *A*-ing only if under the circumstances it would be morally reasonable to expect *S* not to *A*.

Clearly, however, PAE* does not entail PAP, since PAP is a principle about the conditions under which an agent actually is blameworthy and not the conditions under which it would appropriate to hold an agent blameworthy. So if the W-defense is to yield a defense of PAP for blameworthiness, Widerker would first need to show that the fact that it is unreasonable to hold an agent in FR-situations blameworthy is reason to think that the agent is in fact not blameworthy.

This difficulty notwithstanding, some might argue that the W-defense provides intuitive support for the claim that agents in FR-situations should be absolved. After all, perhaps the best explanation of why it would be unreasonable to hold the agent responsible is because he is in fact not responsible. Is there anything the Frankfurt-defender can say in response to this worry? I think there is. I shall offer my own rejoinder to the W-defense shortly, but before doing so I shall first examine three recent replies to the W-defense, which for various reasons I deem problematic.

3 Some Unsuccessful Replies

3.1. *Stump's Reply*

Eleonore Stump (2003) suggests that if we pay careful attention to the way in which the agent's action actually comes about in FR-scenarios and the way it would have come about had the intervener been forced to take action, we will discover an adequate response to the question "what should Jones have done instead?" In an example like Shoot Smith, although it is not up to the agent whether he performs a certain action, it is up to him, Stump says, how the action he performs is produced. More precisely, it is up to him whether his action is a result of his own agency or whether it is the result of Black's intervention. The right answer to Widerker's question, then, is that Jones should have brought it about that his action is a result of Black's intervention and not of his (Jones's) own free agency.⁵

This sort of reply is problematic, however, for a PAP-defender might insist that while perhaps providing a satisfactory response to the question at hand, the reply does so only by pointing to the fact that there are at least some alternative possibilities available to Jones after all, which is precisely what many defenders of PAP have said about FR-situations all along. Stump's response thus avoids the force of the W-defense (if it does) only to fall prey to a different sort of objection to Frankfurt's argument, one according to

⁵ A similar response, which was suggested to me by Al Mele, is that Jones should not have shot Smith freely. What he should have done instead is shoot Smith unfreely; that is, he should have brought it about that he is forced to shoot Smith. Since this is something he could have done, it is a reasonable answer to Widerker's question. The same sorts of worries I raise for Stump's response apply, *mutatis mutandis*, to this response.

which the Frankfurt-style examples do not rule out all morally significant alternative possibilities. This objection has been called the flicker of freedom defense.⁶

Stump anticipates this worry and attempts to forestall it. She says that although the agent does have alternative possibilities of some sort, they are not alternative *actions*. That the agent has control over how his action is produced is not to say that he has control over which action he performs. The agent may have no choice but to perform a certain action while nevertheless retaining control over how that action comes about. PAP-defenders therefore cannot appeal to such alternatives in support of their principle, since PAP is a principle requiring alternative courses of action and not alternative possibilities more generally.

This response seems to me to miss a larger point. Even if what Stump says is correct (and it is, as far as it goes), it would not resolve the larger debate between advocates of the flicker of freedom strategy and Frankfurt-defenders. Advocates of the flicker of freedom defense typically are not committed to PAP *per se* but rather to an account of freedom and responsibility according to which at least some alternative possibilities are required for moral responsibility. PAP provides a natural expression of this view, but should that principle be false, for precisely the sorts of reasons Frankfurt adduces, this would not imply that the alternative possibilities approach to responsibility is mistaken, for as a number of people have pointed out, there may be principles other than PAP which have the consequence that moral responsibility requires alternative possibilities, principles that are not threatened by Frankfurt's case against PAP.⁷ Many of those who have defended Frankfurt's argument, however, including Frankfurt himself, have done so in service of what have come to be known as actual sequence accounts of freedom and responsibility. According to actual sequence theorists, what is important when it comes to moral responsibility is the action the agent actually performed and how that action was brought about. Whatever alternative courses of action, or alternative possibilities more generally, may have been available to the agent are irrelevant to whether the agent is morally responsible. So while Stump's reply might provide the Frankfurt-defender with a technical victory in the case of PAP, her remarks leave the spirit of Frankfurt's argument in doubt insofar as they invite proponents of the flicker defense to press their claim that there are no cases of responsibility without alternative possibilities.

Of course, it may be that the Frankfurt-defender has an adequate reply to the flicker defense. Such a reply, however, would only return us to the controversy over FR and whether a Frankfurt-style example can be constructed that precludes all morally significant alternatives without causally determining the agent's action. Stump's response to the W-defense therefore provides a cogent defense of Frankfurt's argument only if the larger issues surrounding FR have an adequate, Frankfurt-friendly resolution. But part of what I hope to show is that there is a good response to the W-defense that can be given quite independently of the controversy surrounding FR. Moreover, a plausible response to the flicker of freedom defense would likely rule out Stump's reply to the W-defense. In order to respond to the flicker defense, one would have to construct an example in which

⁶ The flicker of freedom strategy is developed in different ways by van Inwagen (1978), Naylor (1984), Otsuka (1998), and Wyma (1997). The term flicker of freedom was first coined by Fischer (1994) as a way of characterizing these sorts of responses to Frankfurt.

⁷ See, e.g., van Inwagen (1978), Otsuka (1998) and Wyma (1997).

any remaining alternatives are either beyond the agent's control or, though within the purview of the agent's control, are morally insignificant in the sense that their presence is irrelevant to the issue of whether the agent is morally responsible. If the remaining alternatives were beyond the agent's control, it would be unreasonable to expect the agent to avail herself of those alternatives. If, on the other hand, the remaining alternatives were morally insignificant, this would seem to contradict the assumption implicit in Stump's remarks that the alternatives to which she claims the agent has access are relevant to whether the agent is morally responsible. In either case, then, Stump's reply to the W-defense would no longer be available.

3.2. Frankfurt, Fischer, and 'Ought' Implies 'Can'

Recently, Frankfurt himself has returned to the fray and has offered a reply to the W-defense. Here is the key part of what he has to say about the matter:

What should Jones have done instead of what he did? He should have behaved virtuously. In deciding of his own free will to [shoot Smith], he violated the requirements of morality. Instead, of doing that, he should have refrained from violating those requirements (Frankfurt, 2003: 343).

Frankfurt's remarks suggest that he is interpreting the question "what should Jones have done?" as a question about what Jones was morally obligated to do. As I explain below, this is apparently not how Widerker intends the question to be understood, but before I do so, it will prove instructive to first examine Frankfurt's answer to the question as he (Frankfurt) seems to be interpreting it.

There are at least two ways to read Frankfurt's remarks.⁸ Frankfurt says that "in deciding of his own free will" to shoot Smith, Jones violated the requirements of morality. This might suggest that on Frankfurt's view Jones violated the demands of morality not by shooting Smith *per se* but rather by shooting Smith *of his own free will*. So what Jones should have done, according to this first reading, is to refrain from shooting Smith of his own free will. The virtue of this reading is that it allows Frankfurt to retain the Maxim that 'ought' implies 'can', since Jones could have refrained from shooting Smith of his own free will by bringing it about that he is forced by Black to shoot Smith. However, this reading of Frankfurt's response is simply a variation of Stump's response and is subject to the same sorts of worries.

A second reading is that Jones violated the demands of morality simply by shooting Smith (full stop). Thus, what he should have done is to refrain from shooting Smith. The virtue of this reading is that it distinguishes Frankfurt's response from that of Stump. The (potential) downside is it also seems to require that Frankfurt abandon the Maxim that 'ought' implies 'can', for on this second reading, Jones should have refrained from shooting Smith even though it was not within his power to refrain from shooting him. Some might find this second reading problematic for precisely this reason, for the Maxim is regarded by many as a highly plausible principle. As Fischer (2006: 203-209) has pointed out, however, the Frankfurt-defender need not find this problematic; it is only natural for those sympathetic with Frankfurt's argument to reject the Maxim along with PAP. Indeed, many actual sequence theorists would point to Frankfurt-style examples as

⁸ I am grateful to Al Mele for pointing me to this first reading.

counterexamples to both PAP and the Maxim. By shooting Smith it seems that Jones has violated an obligation not to murder (or some similar obligation, e.g. not to harm others needlessly), an obligation he would not have if the Maxim were true. Fischer contends, moreover, that rejecting the Maxim “completely disarms the W-defense” (2006: 210). Fischer’s idea seems to be that without the Maxim there is no longer any barrier to saying that Jones should have done otherwise.

There is reason to believe that both Frankfurt (on this second reading) and Fischer have missed their target, for as I noted earlier (in section 2), Widerker insists that the W-defense does not depend on the Maxim. He says there may be circumstances in which an agent is morally obligated to behave in a certain way but in which it would be unreasonable to expect the person to comply with the demands of morality precisely because the person is unable to do so. On his view, then, it appears that Jones may well be morally obligated to refrain from shooting Smith even though it is not within his power to refrain from shooting him. Nevertheless, according to Widerker, it would still be unreasonable to expect Jones to fulfill his obligation not to shoot Smith given that it is an obligation with which he is unable to comply. Denying the Maxim therefore apparently does not disarm the W-defense.⁹

Widerker’s response to Fischer suggests that the question “what should Jones have done?” is not to be understood primarily as a question about what Jones is morally obligated to do but rather as a question about what it would be reasonable to expect him to have done given the circumstances. Thus neither Frankfurt nor Fischer’s reply to the W-defense meets the challenge set out by Widerker, for neither reply provides an answer to Widerker’s question thus construed. An adequate answer to the challenge posed by the W-defender would need to say more than just what the agent was morally obligated to do or not do. It would also need to say what it would be reasonable to expect the agent to do, given the circumstances. If no answer is forthcoming, then, on Widerker’s view, we cannot legitimately hold an agent in a FR-situation blameworthy.

3.3. *The L-Reply*

In my estimation no reasonable answer to the question “what should Jones have done?” can be given. Where I take issue with the W-defense is the assumption that if no reasonable answer is forthcoming, then we should let agents in FR-situations off the hook. As I shall argue shortly, the Frankfurt defender simply has no reason to accept this assumption. However, before doing so, it will prove instructive to look at one last reply to the W-defense on the part of McKenna (2005, 2008).¹⁰

⁹ Fischer (personal correspondence) has suggested to me that, while perhaps not “completely disarming” the W-defense, denying the Maxim might make the W-defense substantially less plausible, for why should one find either PAE or PAE* plausible if one has jettisoned the Maxim? I concede that there might be something to this suggestion, but notice too that there might be reasons for accepting PAE or PAE* even if the Maxim is false. For instance, as I suggest later, in contexts in which one makes judgments about what a person morally ought to do, there is often an accompanying expectation or demand that the person behave as she ought. It is arguably part of our moral practices that we demand of others that they behave as they ought, but such demands seem reasonable only in cases in which people can behave as they ought. Thus, one might find the underlying principles behind the W-defense plausible even if one denies that being morally obligated to act in a certain way implies that one is able to act in that way.

¹⁰ One response to the W-defense that I do not discuss is Zimmerman (2003). He raises some interesting questions about what precisely Widerker means when he speaks of morally reasonable expectations. In the present essay I have been content to rely on an intuitive understanding of this notion.

McKenna concedes that the W-defense cannot be decisively refuted. He and I agree that there is no reasonable answer to Widerker's challenge that does not, in one way or another, play into the PAP-defender's hands. However, McKenna seems to endorse the assumption implicit in the W-defense that holding an agent blameworthy involves an expectation that the agent not behave as she did. It is worth quoting him at some length on this matter:

In a context regarding judgments of blame, in stating what a person should have done but did not do, there is *some* presumption that she could have done that thing.

I think that Widerker's W-Defense cannot be directly refuted; there is something right about it. Widerker has tapped into an important moral intuition favoring a conception of free agency and moral responsibility in terms of alternative possibilities. Some arguing for Frankfurt's thesis are reluctant to admit that there could remain residual intuitions of our moral thought that cannot be fully accommodated. But Widerker's W-Defense has simply hit one of those intuitions spot on. It is only philosophically honest to acknowledge that.... My assessment of the W-Defense is that it is such an impressive argument because it hits upon a raw nerve for the Frankfurt Defender. By way of the W-Defense, certain features of ordinary thought are brought into relief; Frankfurt's conclusions *do* lead to an "irresolvable tension" in our thinking about blame and its implications (regarding what an agent should have done instead) (McKenna, 2008: 784).

Despite his willingness to concede the force of the W-defense, McKenna goes on to argue that its force can be counterbalanced by what he calls the L-reply.

McKenna is an actual sequence theorist. Consequently, he insists that what we should be concerned with is the actual sequence of events. When it comes to freedom and responsibility what matters is not what the agent could or could not have done but what the agent did and how the agent's action was actually produced. So in response to Widerker's question "What should Jones have done instead" McKenna offers the L-reply: "Look at what Jones has done." The L-reply forces us to look not at alternative possibilities or the lack thereof but on the action the agent actually performed and the moral quality of the agent's will that is revealed in that action. By forcing us to look at these features of the situation, the L-reply points to a tension in the thinking of the PAP-defender: those wishing to absolve someone like Jones are forced to deny the intuition that agents who freely and deliberately perform actions they know to be morally wrong, and who thereby display an objectionable quality of will, are blameworthy. In this way McKenna thinks he can defuse the force of the W-defense by showing that, although the W-defender has pointed to an irresolvable tension in Frankfurt-inspired actual-sequence views, the W-defender is committed to a different, though equally irresolvable, tension in his moral thinking when he absolves agents in FR-situations of responsibility. Both views, then, leave us with residual intuitions that we cannot fully explain away. But then the W-defense is not itself sufficient to refute M. Both those who accept and those who

reject M will have to appeal to further considerations to support their respective views about the blameworthiness of agents in FR-situations.¹¹

The virtue of the L-reply is that once we focus on the actual sequence of events and the moral quality of the agent's will that is revealed in his action, it does seem quite plausible to suppose that agents in FR-situations who perform morally wrong actions are blameworthy for what they have done. The difficulty with McKenna's reply, as I see it, is that it concedes too much to the W-defender. Those wishing to defend Frankfurt's thesis, I shall argue, need not admit that their position leaves them with an "irresolvable tension" in their thinking about blame and its implications. I think further examination of the W-defense will reveal that the moral intuition which that argument "taps into," as McKenna puts it, is not as plausible as it might first appear.

4 A Rejoinder to the W-Defense

Having explained why I find some of the previously published responses to the W-defense unsatisfying, I now develop a rejoinder of my own. As we have seen, it is the W-defender's contention that holding a person blameworthy for her wrongdoing essentially involves an expectation or demand on our part that the person not have behaved as she did. But the W-defender rightly insists that such an expectation would be unreasonable if the agent was unable to avoid acting as she did. Thus, in situations in which a person could not have avoided performing some action, we cannot reasonably blame that person for what she has done. This argument can be captured in the following schema:

- (1) To hold *S* blameworthy for *A*-ing essentially involves an expectation that *S* not have *A*-ed.
- (2) But to expect that *S* not have *A*-ed would be unreasonable if *S* could not have done otherwise than *A*.
- (3) Hence, holding *S* blameworthy for *A*-ing when *S* could not have done otherwise than *A* would involve an unreasonable expectation on the part of those holding *S* blameworthy.

I take it that (3) constitutes the sort of "irresolvable tension" in Frankfurt's thought that McKenna thinks the W-defense brings into relief. It points to an apparent unreasonableness to which one would have to be committed if one were to embrace Frankfurt's conclusions about the blameworthiness of agents in FR-situations.

How cogent is the W-defense thus construed? Although (2) looks quite plausible, why should the Frankfurt-defender concede (1)? Why think that blaming a person for what she has done essentially involves an expectation or demand that she not have done what it is we are blaming her for doing? The answer is not at all obvious.

Perhaps one reason some might find (1) appealing has to do with the fact that in paradigm cases of blameworthiness it is plausible to suppose that the agent could have done otherwise. It is, I think, part of our commonsense picture of ourselves as free agents that we can often do otherwise than what we in fact do. So in normal cases in which it is

¹¹ For instance, the Frankfurt-defender might appeal to the two arguments (mentioned in section 1) given by Frankfurt in support of M. If these arguments could be defended, then the force of the W-defense might well be outweighed by these arguments together with L-reply. For a defense of Frankfurt's further arguments regarding M and how they can be used to counteract the W-defense see McKenna (2008).

appropriate to blame a person for her actions, it might well be reasonable to expect that the person have done otherwise. It might even be the case that in most instances there *is* an expectation that the person have done otherwise. But clearly this need not imply that all instances of holding someone blameworthy must involve an expectation that the person have acted differently. Perhaps there are non-paradigmatic cases in which we may reasonably blame someone for what she has done even though we do not (and could not reasonably) expect her to do otherwise.

In his original essay on PAP, Frankfurt pointed out that that principle gains plausibility from the fact that when considering cases in which a person could not have done otherwise, we tend to think only of those cases in which the circumstances that prevent the person from doing otherwise also play a role in producing the agent's action. Frankfurt suggests, however, that if we focus on atypical, non-paradigmatic cases, namely, FR-situations, it becomes significantly less clear that PAP is true. Focusing only on the paradigmatic instances of "couldn't have done otherwise," cases in which the agent is forced, coerced, or otherwise compelled against her will, does indeed suggest something like PAP. It is only by looking at the non-paradigmatic cases, however, that we can discover potential counterexamples to PAP.

A similar point applies to (1). If we focus only on paradigmatic cases of holding blameworthy, it might seem plausible to suppose that all cases of holding blameworthy involve an expectation that the person have done otherwise, for as I have suggested, in those paradigmatic cases we may typically think it is reasonable to expect the person to have done otherwise—because we believe the person could have done otherwise—and we may in fact expect the person to have done otherwise. But in order to assess the plausibility of (1), we need to look at atypical cases as well. Once we do so, however, it is less clear that holding someone blameworthy essentially involves an expectation that the person have acted differently. As McKenna rightly points out, when we focus on the actual sequence of events in FR-situations, and especially the fact that agents in those situations act entirely on their own accord, there is a good deal of pressure to say that the agent is blameworthy despite the fact that we could not reasonably expect the person to do otherwise. So in the absence of further reason to think that holding a person blameworthy essentially involves an expectation that the person have done otherwise, those who share these intuitions about the blameworthiness of agents in FR-situations are well within their rights to deny (1) in favor of M.

A second reason some might find (1) initially plausible has to do with the fact that holding a person blameworthy invariably involves a conative element. When we blame someone for what she has done, we desire that the person not have done what she did. This desire, moreover, is arguably essential to holding a person blameworthy. To see this, imagine that the devil is evaluating Jones's behavior. The devil might agree with Frankfurt at the cognitive level that Jones is *blame-worthy*—i.e. that he meets the relevant control and epistemic conditions for moral responsibility—but such a cognitive judgment on the part of the devil does not mean that the devil actually *holds* Jones blameworthy. After all, the devil is no doubt gleeful that Jones has done the morally wrong thing. Indeed, we may suppose that it was the devil who sent Black to insure that Jones shoot Smith! In order for the devil to actually hold Jones blameworthy, there needs

to be a conative element present as well. In order to actually blame Jones, the devil must also desire that Jones not have shot Smith.¹²

So blame involves both a cognitive and a conative element, and it is perhaps not difficult to see how a desire that *S* not *A*, which is essential to holding *S* blameworthy for *A*-ing, might be confused with, or seen as involving or somehow being associated with, an expectation or demand that *S* not *A*. Typically, when we expect a person not to do something, our expectation is accompanied by a desire that the person not do that thing. It is important to notice, however, that to desire that *S* not *A* needn't point to the presence of an expectation or demand that *S* not *A*. I might desire that you perform a certain action even if I cannot reasonably expect you to perform that action. I might desire that you give me a million dollars, for example, but it would clearly be unreasonable for me to demand this of you. Or, to return to FR-situations, we might desire that Jones not shoot Smith, but given the circumstances it would be unreasonable to expect Jones not to shoot Smith. The fact that holding a person blameworthy involves a conative element therefore does not support (1) either.

Widerker claims that an expectation that a person not act as she did is essential to our moral disapproval of her behavior. The Frankfurt-defender, I am suggesting, has been given no reason to accept this claim. Why must we expect a person to do otherwise in order to disapprove morally of the person's behavior? It seems entirely possible to disapprove of a person's behavior and indeed to hold that person blameworthy for what she has done without there being an expectation on our part that she not have behaved in that way. To be sure, we will no doubt wish that the person not have behaved as she did, and we may judge that the person ought not have behaved in that way, but as I have argued, this needn't involve an expectation or demand on our part that the person not behave as she did.

5 Conclusion

So what should Jones have done instead of shooting Smith? Well, if we mean "what *ought* he to have done instead?" then it is open to the Frankfurt-defender to deny that 'ought' implies 'can' and to insist that Jones should have refrained from shooting Smith. If, on the other hand, we mean "what would it be reasonable to expect (or demand) Jones to have done instead?" then the correct answer, I think, is "there is nothing we could have reasonably expected him to do instead." Since Jones could not have done otherwise than shoot Smith, it would be quite unreasonable to expect him not to shoot Smith. I have argued that this is problematic for the Frankfurt-defender only if by holding Jones blameworthy the Frankfurt-defender is thereby committed to expecting that Jones not have done what he did. But there is no reason to suppose the Frankfurt-defender must be committed to such an expectation, for there is no reason to suppose that an expectation that a person not have acted in a certain way is essential to our moral disapproval of that person's behavior or to our holding that person blameworthy for what she has done.

I conclude with a tentative observation. Up to this point, the debate over PAP has largely focused on metaphysical considerations regarding the nature of free agency. However, my discussion of the W-defense suggests that many of the intuitions driving various positions in this debate might depend to a greater or lesser extent on particular

¹² I owe this point about the conative element involved in blame and the corresponding example to [suppressed].

accounts of the nature of moral responsibility and allied notions such as praise and blame. In light of this, the debate over PAP might benefit from a closer examination of these moral concepts. Whether moral responsibility requires the ability to do otherwise might depend, in the final analysis, on which account of moral responsibility one adopts. In this paper I have argued that the account of blame implicit in the W-defense has little to recommend itself and seems implausible in light of intuitions arising from Frankfurt-style examples, which leads me to the conclusion that the W-defense is ultimately unsuccessful.

References

Fischer, J. M. (1994). *The metaphysics of free will: an essay on control*. Oxford: Blackwell Publishers.

Fischer, J. M. (2006). *My way*. New York: Oxford University Press.

Frankfurt, H. (2003). Some thoughts concerning PAP. In Widerker and McKenna (2003): 339-48.

Frankfurt, H. (1969). Alternate possibilities and moral responsibility. *Journal of Philosophy* 66: 829-39.

Ginet, C. (1996). In defense of the principle of alternative possibilities: why I don't find Frankfurt's argument convincing. *Philosophical Perspectives* 10: 403-17. Reprinted with an addendum in Widerker and McKenna (2003): 75-90.

Hunt, D. (2005). Moral responsibility and buffered alternatives. *Midwest Studies in Philosophy* 29: 126-45.

Hunt, D. (2000). Moral responsibility and unavoidable action. *Philosophical Studies* 97: 195-227.

Kane, R. (1985). *Free will and values*. Albany, NY: State University of New York Press.

McKenna, M. (2003). Robustness, control, and the demand for morally significant alternatives. In Widerker and McKenna (2003): 201-217.

McKenna, M. (2005). Where Strawson and Frankfurt meet. *Midwest Studies in Philosophy* 29: 163-80.

McKenna, M. (2008). Frankfurt's argument against the principle of alternative possibilities: looking beyond the examples. *Nous* 42: 770-793.

Mele, A. and David Robb. (1998). Rescuing Frankfurt-style cases. *Philosophical Review* 107: 97-112.

Naylor, M. (1984). Frankfurt on the principle of alternate possibilities. *Philosophical*

Studies 46: 249-58.

Otsuka, M. (1998). Incompatibilism and the avoidability of blame.” *Ethics* 108: 685-701.

Pereboom, D. (2000). Alternate possibilities and causal histories. *Philosophical Perspectives* 14: 119-38.

Stump, E. (2003). Moral responsibility without alternative possibilities.” In Widerker and McKenna (2003): 139-158.

van Inwagen, P. (1978) “Ability and Responsibility.” *Philosophical Review* 87: 201-24.

Widerker, D. (1995). Libertarianism and Frankfurt’s attack on the principle of alternative possibilities. *Philosophical Review* 104: 247-61.

Widerker, D. (2000). “Frankfurt’s attack on alternative possibilities: a further look. *Philosophical Perspectives* 14: 181-201.

Widerker, D. (2003) “Blameworthiness, and Frankfurt’s Argument Against the Principle of Alternative Possibilities.” In Widerker and McKenna, eds. (2003) 53-74.

Widerker, D. (2005). Blameworthiness, non-robust alternatives, and the principle of alternative possibilities. *Midwest Studies in Philosophy* 29: 292-306.

Widerker, D. and Michael McKenna, eds. (2003). *Moral responsibility and alternative possibilities*. Aldershot, UK: Ashgate Press.

Wyma, K. (1997). Moral responsibility and leeway for action. *American Philosophical Quarterly* 34: 57-70.

Zimmerman, M. J. (2003). The moral significance of alternate possibilities. In Widerker and McKenna (2003): 301-26.