

Journal of Cognition and Neuroethics

Rescuing PAP from Widerker's Brain-Malfunction Case

Greg Janzen

Osgoode Hall Law School

Biography

Greg Janzen has been a lecturer in philosophy at Mount Royal University and the University of Calgary. Some of his work has appeared in *Journal of Consciousness Studies*, *Philosophia*, *Journal of Philosophical Research*, and *Erkenntnis*. He is also the author of *The Reflexive Nature of Consciousness* (John Benjamins Publishing, 2008). He is currently attending Osgoode Hall Law School in Toronto.

Publication Details

Journal of Cognition and Neuroethics (ISSN: 2166-5087). September, 2015. Volume 3, Issue 2.

Citation

Janzen, Greg. 2015. "Rescuing PAP from Widerker's Brain-Malfunction Case." *Journal of Cognition and Neuroethics* 3 (2): 1–22.

Rescuing PAP from Widerker's Brain-Malfunction Case

Greg Janzen

Abstract

According to the principle of alternate possibilities (PAP), a person is morally responsible for what she has done only if she could have done otherwise. David Widerker, a prominent and long-time defender of this principle against Harry Frankfurt's famous attack on it, has recently had an unexpected about-face: PAP, Widerker now contends, is (probably) false. His rejection of PAP is a result, in large part, of his coming to believe that there are conceptually possible scenarios, what he calls 'IRR-situations,' in which circumstances that nowise bring it about that an agent performs a particular action are precisely the circumstances that make it impossible for her to avoid performing that action. The circumstances that guarantee that the agent will perform the action turn out to be immaterial, since the agent in an IRR-situation is blameworthy because she would have performed the action even if, contrary to the specified facts, an alternative course of action had been available to her. The goal of this article is to show that Widerker's report of PAP's demise has been an exaggeration: careful scrutiny reveals that the kind of scenario that he believes refutes the principle—his 'brain-malfunction' scenario—is not an IRR-situation at all.

Keywords

Principle of Alternate Possibilities, Free Will, Moral Responsibility, Harry Frankfurt, Frankfurt Cases, Compatibilism, David Widerker

Introduction

According to the principle of alternate possibilities (PAP), a person is morally responsible for what she has done only if she could have done otherwise.¹ This principle

-
1. So stated, PAP is equivalent to Frankfurt's original formulation, which says that 'a person is morally responsible for what he has done only if he could have done otherwise' (1969, 829). But this formulation admits of at least two interpretations: (i) as expressing the proposition that a person is morally responsible for performing a particular action, *V*, only if she could have *avoided* (or *refrained* from) doing *V*, or (ii) as expressing the proposition that a person is morally responsible for *V* only if she could have avoided doing *V* and, in addition, could have performed *some other action*. This latter interpretation is very non-standard—neither friends nor critics of PAP interpret it thus—and there is no reason to interpret it thus. Accordingly, PAP, in what follows, should be taken as expressing the proposition that an agent is morally responsible for doing *V* only if he could have avoided doing *V*. A further bookkeeping point is that, unless otherwise noted, 'moral responsibility' should be regarded as synonymous with 'morally blameworthy,' since Widerker's project (in his 2006) is to develop an account of moral blameworthiness. Compare his

is indisputably philosophically important, frequently playing a pivotal role in debates regarding the sort of freedom required for moral responsibility. David Widerker, a prominent, eloquent, and long-time defender of the principle against Harry Frankfurt's (1969) famous attack on it (Widerker 1995; 2000; 2003; 2005), has recently had an unexpected about-face. PAP, Widerker now contends, is (probably) false: an agent can be morally responsible for what she has done even though she could not have done otherwise.² He now subscribes to a view that he calls 'Frankfurt-friendly libertarianism,' which is a libertarian-based account of freedom and moral blameworthiness that rejects PAP (Widerker 2006, 2009). Widerker's rejection of PAP is a consequence, in large part, of his coming to believe that there are conceptually possible scenarios, what he calls 'IRR-scenarios' or 'IRR-situations,' in which circumstances that nowise bring it about that an agent performs a particular action are precisely the circumstances that make it impossible for her to avoid performing it. The circumstances that guarantee that the agent will perform the action turn out to be immaterial, since the agent in an IRR-situation is blameworthy because she would have performed the action even if, contrary to the specified facts, an alternative course of action had been available to her.³ The goal of this article is to show that Widerker's report of PAP's demise has been an exaggeration: careful scrutiny reveals that the kind of scenario he believes refutes the principle is not an IRR-situation at all.

1. Widerker's rejection of PAP

Let me begin by explaining how the axis of Widerker's examination regarding PAP became rotated so that he now rejects the principle.

Widerker (2006, 168) characterizes an IRR-situation as follows (where a state of affairs is *causally possible* at time *t* if and only if it obtains in some possible world that up until, and including, *t* has the same history and the same laws of nature as the actual

rendering of PAP: 'An agent *S* is morally blameworthy for performing a given act *V* only if he could have avoided performing it' (2006, 163).

2. This is not to say that Widerker now believes that PAP lacks intuitive appeal. On the contrary, he still considers it very intuitively appealing, since it is accords with 'the intuitive link between attributing blame to an agent and expecting of the agent that he not have done what he did' (Widerker 2009, 90).
3. It is not clear to me why Widerker calls them IRR-situations, but my guess is that it is because the circumstances that close off alternative possibilities are *irrelevant* to whether the agent performs the action in question.

world, and where an alternative is *actionally accessible* to an agent at t if and only if the agent has it within her power to bring about that alternative at t):

- a. At t , S decides-to- V on his own.
- b. P_1, P_2, \dots, P_n are *all* (in the circumstances) causally possible alternatives to S 's deciding-to- V at t .
- c. Each P_i ($i = 1 - n$) is *actionally inaccessible* to S in the circumstances.

This characterization of an IRR-situation, Widerker claims, is in keeping with Frankfurt's (1969, 830, 837) original characterization as a situation in which the agent's decision, though unavoidable, is in no way brought about by the circumstances in which it occurs.⁴ Widerker now believes that IRR-situations are conceptually possible, and he has developed an example of such a situation himself.

What is the significance of IRR-situations, assuming that they are conceptually possible? Do they imply that PAP is false? In the past, Widerker was reluctant to accept this conclusion, because if PAP is false, one has to assume that an agent in an IRR-situation who knowingly performed a wrong action is blameworthy despite being unable to avoid performing it. But Widerker no longer finds this assumption objectionable. 'I find it intuitive,' he remarks, 'that in [certain IRR-situations] the agent is blameworthy for the decision he made, even though it was not within his power not to make it' (2009, 91). By Widerker's lights, then, PAP has been falsified, since the following two conditions for falsifying the principle have been satisfied: (1) an IRR-situation has been described, and (2) the agent in the IRR-situation is blameworthy for his action.

By way of introducing Widerker's example of an IRR-situation, consider the following purported Frankfurt-style counterexample to PAP that, on his view, does *not* describe an IRR-situation:

4. Widerker also provides a more general notion of an IRR-situation, which he calls a 'G-IRR-situation'. In a G-IRR-situation, the agent lacks a *morally significant alternative*, where V^* is a morally significant alternative to a morally wrong act V that an agent S performs at t if and only if: '(i) V and V^* are incompatible, (ii) S is aware (or should be aware) that if he performs V^* at t , he will not be acting in a morally wrong way, (iii) S truly believes that it is within his power to perform V^* at t , and (iv) in the circumstances, V^* -ing at t would be regarded as a reasonable way for S to avoid V -ing at t ' (2006, 175). As Widerker points out, since the notion of a G-IRR-situation is more general than that of an IRR-situation, constructing an example of a G-IRR-situation should be easier than constructing an example of an IRR-situation. For the time being, I will disregard the distinction between IRR-situations and G-IRR-situations and continue to speak only about the former. The notion of a morally significant alternative will remerge, however, in my argument (in the next section) that Widerker has not constructed an IRR-situation.

PROMISE-BREAKING (PB): Jones is deliberating as to whether to keep the promise he made to his uncle to visit him in the hospital shortly before a critical operation his uncle is about to undergo. Jones is his uncle's only relative, and the visit is very important to the uncle. The reason for Jones's deliberating is that, on his way to the hospital, he (incidentally) met Mary—a woman with whom he was romantically involved in his distant past and whom he has not seen since then. Mary, being eager to talk to Jones, invites him for a cup of coffee in a nearby restaurant. She explains that she is in town just for a couple of hours, and wishes to spend those hours with him. Jones is aware that if he accepts Mary's offer, he will not be able to make it to the hospital during visiting hours. Unbeknownst to Jones, there is another person, Black, who for some reason does not want Jones to visit his uncle. Black has the power and the means to force Jones to decide to stay with Mary. 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 that he is going to decide to break his promise to his uncle. Call that sign Q. If Jones shows that sign, then Black does nothing, knowing that in this case Jones will decide to accept Mary's invitation. (It is assumed that Black knows Jones very well in this regard.) Finally, suppose that Black does not have to intervene, since Jones decides on his own not to keep the promise, so as to be able to spend time with Mary. (Widerker 2006, 164)

Here it would seem that Jones is blameworthy for breaking his promise because Black exerted no influence on his behaviour (even though he made it impossible for him to avoid breaking it).

According to Widerker (2006, 165–6), however, PB founders on the dilemma objection.⁵ Exponents of this objection advance the following dilemma regarding the relation between Q, absence of which would have led Black to intervene, and Jones's decision. Either Q (or a condition whose presence is indicated by Q) determined Jones's decision, given the circumstances, or it did not. If it did, then the decision was brought

5. This objection, which has been suggested by a number of philosophers, most notably Widerker (1995), Ginet (1996), and Kane (1996, 142–4, 191–2), has spawned a small sub-industry in the literature on the debate over Frankfurt-style counterexamples. For important critical responses to it, see Mele and Robb (1998), Fischer (1999; 2010), McKenna (2003), and Haji and McKenna (2004).

about by the conditions in which it occurred and thus we do not have an IRR situation. If, on the other hand, Q was merely a reliable indicator of Jones's decision, then he might have exhibited Q, which would have prompted Black to refrain from intervening, but decided to keep his promise anyway. On the first horn of the dilemma, Jones's decision was determined to take place given the conditions in which it was made, and on the second horn, he could have decided otherwise. Either way, PB is not an IRR-situation.

Widerker, though, now believes that the dilemma objection fails,⁶ for there are scenarios that, unlike PB, *do* describe an IRR-situation and that are not vulnerable to dilemma worries. The following variant of PB, he claims, describes such a situation:⁷

BRAIN-MALFUNCTION (BMF): As in Promise Breaking, Jones deliberates as to whether to accept Mary's offer, and ultimately decides on his own at *t* to violate the promise to his uncle. Normally, one can avoid deciding as one does by deciding otherwise. But in our scenario Jones does not have that option, since shortly before beginning to deliberate, he undergoes a neurological change as a result of which one of the (neurological) causally necessary conditions for his deciding otherwise, a condition which we may call *N*, does not obtain. Let us also assume that all this is unknown to Jones (who believes that he can decide to keep the promise), and that *N*'s absence does not affect his deliberation process. Note that these assumptions do not render Jones's actual decision of not keeping the promise ($D(-K)$) causally determined. The fact that *N* is a causally necessary condition of Jones's deciding to keep the promise entails that *N*'s absence is sufficient for its not being the case that Jones decides to keep it, that is, for $-D(K)$. But from this it does not follow that *N*'s absence is sufficient for $D(-K)$ [W]e have not yet eliminated all the alternatives to Jones's deciding not to keep the promise that are accessible to him. There is the alternative of his continuing to deliberate at *t*, instead of making $D(-K)$. However, this alternative can be ruled out either by stipulation or by assuming that in the scenario at hand Jones needs to make a

6. Though he still defends it against certain criticisms: see Widerker and Goetz (2013).

7. See McKenna (2003, 209–10) for a similar example, which, as it happens, he developed jointly with Widerker. My argument that Widerker's example is not an IRR-situation applies, *mutatis mutandis*, to McKenna's.

decision right away, as otherwise he will miss the bus that can get him to the hospital on time. If that is the case, then in our scenario, the option of continuing to deliberate is practically equivalent to deciding not to keep the promise. (Widerker 2006, 169–70)

This scenario is not vitiated, it would seem, by the dilemma objection, since there is no sign Q that could generate the dilemma between lack of libertarian freedom, on the one hand, and the presence of an actionally accessible alternative, on the other. Thus, on Widerker's view, BMF satisfies the criteria for an IRR-situation: 'In it, the only alternatives to [Jones's deciding not to keep his promise] are alternatives which, though causally possible, are actionally inaccessible to Jones' (2006, 171).

But is BMF an IRR-situation? In the next section I will attempt to show that it is not, and hence that it is not deleterious to PAP. As we have seen, an IRR-situation must feature an agent who (1) is blameworthy for his decision and (2) lacks an actionally accessible alternative. I shall argue that BMF fails to meet this second requirement.⁸

2. BMF is not an IRR-situation

The reason that BMF is not an IRR-situation is that Jones has an actionally accessible alternative: instead of deciding to break his promise at *t*, he can continue deliberating (in a way shortly to be elucidated).⁹ Widerker, as we have seen, rejects this alternative on the ground that it is not actionally accessible. He says, to repeat, that:

[the alternative of continuing to deliberate] can be ruled out either by stipulation or by assuming that in the scenario at hand Jones needs to make a decision right away, as otherwise he will miss the bus that can get him to the hospital on time. If that is the case, then in our

8. I am not the first person to defend PAP against Widerker's attack on the principle. At least one other commentator, Carlos Moya (2007), has also taken up the challenge, arguing that BMF fails to meet the first requirement. For a cogent reply to Moya, see Widerker (2009). (In section 2.2 I briefly discuss Moya's argument and Widerker's rejoinder.) Widerker (2006, 169) has proposed another example of an IRR-situation, which he calls '*Z-persons*,' but I will not discuss that example here. Although Moya's argument that BMF fails to meet the first requirement for an IRR-situation does not refute BMF, it does, it seems to me, refute *Z-persons*.

9. Note that since it is inevitable that Jones will fail to keep his promise after his neurology fails him and he begins deliberating, the claim being disputed is that he has no alternative to *deciding* to break it. This decision, I contend, is not one that he has to make.

scenario, the option of continuing to deliberate is practically equivalent to deciding not to keep the promise. (2006, 170)

But there is ample room for disagreement here. My strategy will be to show that the alternative of continuing to deliberate cannot be foreclosed in the ways Widerker describes. I first, in section 2.1, argue that it cannot be ruled out by adding a time constraint. Then, in section 2.2, I argue that it cannot be ruled out by stipulation.

2.1 The alternative of continuing to deliberate cannot be ruled out by adding a time constraint

According to Widerker, once we assume that Jones needs to make a decision right away (as otherwise he will, for example, miss the bus that can get him to the hospital on time), ‘the option of continuing to deliberate is practically equivalent to deciding not to keep the promise’. But is it?

There are two ways to interpret this equivalence claim: (i) as the claim that continuing to deliberate is *morally* equivalent to deciding to break the promise, or (ii) as the claim that continuing to deliberate is *one and the same* (mental) action as deciding to break the promise.¹⁰ I will examine these interpretations in turn.

On this latter interpretation, the equivalence claim seems clearly false. If Jones continues deliberating—say because he is conflicted about what to do and is unable to make a decision—then he has not made a decision at all. Quite the reverse, he has been *indecisive*, a not uncommon phenomenon even among morally competent agents, and his indecisiveness has resulted in his breaking his promise. It is true that Jones *knows* (at some level) that if he does not make a decision by *t*, then, at *t*, he will break his promise, but it does not follow from this that if he fails to make a decision by *t*, then, at *t*, he has decided to break his promise. It is more accurate to say that Jones’s breaking his promise is a *foreseeable consequence* of his continuing to deliberate, not that his continuing to deliberate is the same action as his deciding to break his promise.

It merits emphasizing that Widerker understands ‘action’ in Ginet’s (1990) sense, i.e., as ‘either a causally simple mental action such as a volition-to-*V* or the forming of

10. An anonymous reviewer has suggested a third interpretation of the equivalence claim: that continuing to deliberate is *tantamount* to deciding to break the promise, without being either morally or numerically equivalent to breaking it. If, however, the former is tantamount to the latter, then the former is equivalent in seriousness to the latter, so it would seem that on a ‘tantamount’ interpretation of the equivalence claim, continuing to deliberate is morally equivalent to deciding to break the promise. In short, the tantamount interpretation collapses into the moral equivalence interpretation.

an intention-to-*V*, or a complex action such as Sam's action of killing Smith that consists of a simple mental action causing the event of Smith's death' (Widerker 2006, 176n26). Notice that, in continuing to deliberate, Jones has neither a volition nor an intention to break his promise. Indeed, it is precisely the *lack* of those elements in his practical reasoning process that makes it the case that his continuing to deliberate is not the same action as his deciding to break his promise. Thus, on Widerker's own understanding of what constitutes an action, Jones's continuing to deliberate is not the same action as his deciding to break his promise.

The equivalence claim, then, is better understood as the claim that the act of continuing to deliberate is morally equivalent to deciding to break the promise. And, in fact, in his 2009 restatement of his 2006 argument for the conceptual possibility of IRR-situations, Widerker states the equivalence claim using moral equivalence terminology. '[T]he act of continuing to deliberate,' he says, 'is morally equivalent to the decision not to keep the promise, and hence does not count as a *morally significant alternative*' (2009, 90).¹¹ The idea is that, even if continuing to deliberate and deciding to break the promise are not the same action, continuing to deliberate is not an actionally accessible alternative, since even if Jones availed himself of this option, he would not avoid blame.

But this is eminently disputable. Following Widerker, I will call an alternative that is actionally accessible in the moral sense a 'morally significant alternative'. Despite initial appearances, Jones's continuing to deliberate is such an alternative. To see this, consider the following counterfactually modified version of BMF in which Jones undergoes a moral struggle as a result of the absence of *N*:

BRAIN-MALFUNCTION* (BMF*): Jones, while deliberating in a normal fashion about whether to keep his promise, adopts a favourable attitude toward keeping it and, as a result, forms a desire to keep it. At this point, due to the absence of *N*, he hits a wall (so to speak). His neurological deficit prevents him from acting on his desire and deciding to keep it. His desire to keep it persists, however, and he verges on so deciding. At this point, *N*'s absence again prevents him from acting on his desire, so he continues deliberating. This back-and-forth goes on until the moment he runs out of time to deliberate.¹² At the last

11. Widerker, as will be recalled, initially claimed that 'the option of continuing to deliberate is *practically* equivalent to deciding not to keep the promise' (2006, 170, emphasis added).

12. Cf. Larvor (2010, 507).

second, he once again develops a desire to keep his promise, verges on deciding to keep it, but *N*'s absence stymies his 'effort' to do the right thing. In the end, he runs out of time to deliberate, misses the bus to the hospital, and thereby breaks his promise.

With an eye toward forestalling the objection that I am begging the question against Widerker's stipulation that Jones (in BMF) is unable to decide to keep his promise, it is worth making explicit two implicit assumptions on which I proceed in this case: (i) that one can adopt a favourable attitude toward doing *V* without having decided to do *V*, and (ii) that one can have a desire to do *V* without having decided to do *V*—for example, I can have a desire to go on holiday without having decided to go on holiday. Although Jones, then, in BMF*, adopts a favourable attitude toward keeping his promise and forms a desire to keep it, he does not decide to keep it.

Now let us address the question whether Jones is blameworthy in the (counterfactual) scenario described in BMF*. Intuitions, while not inviolable, are an indispensable starting point, and I submit that, intuitively, he is not blameworthy. Had he not been frustrated by circumstances beyond his control—i.e., by the absence of *N*—he would have decided to keep his promise. Indeed, but for the absence of *N*, he would have stopped deliberating when he formed the desire to keep it and, at that point, decided to keep it. To be sure, moral responsibility is not an all-or-nothing affair: the moral disvalue (or value) of an action, and hence the degree of blameworthiness (or praise) we attribute to the agent who performed the action, is usually a matter of degree.¹³ Since (we may suppose) he momentarily contemplated breaking his promise in the first place, therefore, Jones is, arguably, *somewhat* blameworthy in the counterfactual sequence in which he continues deliberating. But if he is merely somewhat blameworthy, then his continuing to deliberate is a morally significant alternative; and this is because had he simply decided outright to break his promise, he would have been *wholly* blameworthy. Continuing to deliberate, then, mitigates his blameworthiness (even if it does not exempt him from it altogether).

It bears underscoring, however, that we cannot rule out the possibility that Jones never so much as entertains the idea of breaking his promise. More fully, we cannot rule out the possibility that his entire deliberative process (prior to *t*) consists of his repeatedly verging on deciding outright to keep it (only to be foiled, time and again, by the absence of *N*). In other words, Jones has an alternative that exempts him from blame altogether.

13. I discuss this further in section 3 below.

Friends of PAP, in any event, will insist that we cannot rule out this possibility, at least not without begging the question against PAP. Either way—whether he is somewhat blameworthy or not blameworthy at all—Jones’s continuing to deliberate, in the way described in BMF*, is *not* morally equivalent to his deciding to break his promise.

We can augment the case for this conclusion by formulating a converse scenario in which Jones is unable to decide to *break* his promise. So suppose that, shortly before beginning to deliberate, he undergoes a neurological change as a result of which one of the neurological causally necessary conditions for his deciding to break his promise does not obtain. Further suppose that, while deliberating, he adopts a favourable attitude toward breaking it and, as a result, forms a desire to break it. Finally, suppose that a back-and-forth similar to the one described in BMF* transpires, such that, in the end, Jones runs out of time to deliberate and thereby keeps his promise. Here Jones is not praiseworthy because he continued deliberating despite wanting to break his promise. But this situation parallels the situation described in BMF*, just in converse, so if Jones is not praiseworthy in this scenario, then, by parity of reasoning, he is not blameworthy in BMF*.¹⁴

Note that it does not matter—and with this Widerker would agree—that Jones’s acts as described in BMF* occur covertly within consciousness. Mental acts, such as deciding, concentrating, pondering, undertaking, imagining, forming an intention, mulling over, etc., are still acts, things we do. Moreover, for libertarians, mental acts constitute the *loci* of moral responsibility, so any non-question begging argument against PAP cannot appeal to the distinction between mental acts and non-mental acts.

The following principle bolsters the intuition that Jones is not blameworthy in BMF* (where ‘*V*’ is an action that *S* is obligated to perform at *t*):

(NB) If, through no fault of her own, *S* is prevented from doing *V* despite doing everything in her power to do *V*,¹⁵ then *S* is not blameworthy for failing to do *V*.¹⁶

14. There are asymmetries between moral praise and moral blame (for discussion, see Widerker 1991), but, as far as I can tell, they do not affect the point made here.

15. Does it matter that, in BMF*, Jones is not *aware* that he is doing everything in his power to do *V*? I don’t see that it does. In many circumstances, to be sure, being absolved of blame (or being held blameworthy, or whatever) will no doubt depend on one’s being aware of one’s actions, but in the circumstances under consideration it does not.

16. The ‘through no fault of her own’ clause is important. If *S* is antecedently responsible for bringing about whatever prevented her from being able to do *V*, then arguably she is blameworthy for failing to do *V* even

This principle is reasonable. Indeed, it is not clear how one might argue for it, beyond simply adducing examples involving agents who, by doing everything in their power to perform obligatory actions, avoid blame for failing (through no fault of their own) to perform them. But NB is itself undergirded by the celebrated Kantian maxim that ‘ought’ implies ‘can’. Why does an agent, by doing everything in her power to perform an obligatory action, avoid blame for failing (through no fault of her own) to perform it? A natural answer is that she avoids blame because she *cannot* perform the action in question: circumstances beyond her control have rendered her powerless to do what she is obligated to do. Thus, the principle that ‘ought’ implies ‘can’ supports NB.¹⁷

At the risk of being slightly tedious, I will state the main argument of this section in another way. On Widerker’s view, Jones is unable to avoid deciding to break his promise. Conditions are such that either he will decide to break it on his own or he will run out of time to deliberate (and thereby do something morally equivalent to deciding to break it). There is no third possibility, at least none that Jones can actualize by his own efforts and that will exempt him from blame. I have advanced considerations suggesting that this is false, that there is a third possibility, one that Jones can actualize by his own efforts and that will exempt him from blame: he can continue deliberating in such a way that he adopts a favourable attitude toward keeping his promise and, as a result, forms a desire to keep it. If this desire persists, or keeps recurring, such that he continues deliberating until the moment he runs out of time to deliberate, then he has avoided blame, since he has done everything in his power to decide to keep his promise (only to be thwarted, unbeknownst to him, by a neurological deficit). We have, then, a morally significant alternative: Jones continues deliberating in such a way that he avoids blame.

Of course, the ‘in such a way’ clause here is important. If Jones continues deliberating for purely selfish reasons—if, for example, he continues deliberating solely because he is worried that his uncle will not bequeath to him his valuable coin collection if he breaks his promise—then his continuing to deliberate does not exempt him from responsibility.¹⁸

if she did everything in her power to do *V*.

17. Kant’s maxim has been contested, of course, but it is intuitively appealing, theoretically and explanatorily important—e.g., it explains why it makes no sense to say of the wheelchair bound person that he ought to wade into the pond to save the drowning child—and sufficiently widely accepted to justify assuming it here without further argument.
18. Would deliberating in this way exempt him from responsibility for breaking his promise (even if it does not exempt him from responsibility for, well, deliberating in this way, which is itself objectionable)? Perhaps, but not obviously. Suppose that his uncle chastises him for breaking his promise. Intuitively, it will not do

Jones avoids blame only if he continues deliberating in the *appropriate* way, i.e., in a way that exhibits a proper respect for morality (like the way described in BMF*).

To be sure, in the scenario, BMF, that Widerker initially invites us to imagine, Jones simply decides, on his own, to break his promise without forming a desire to keep it. But that is irrelevant, since Jones, if he is libertarianly free, *has it within his power* to perform (mental) actions that exempt him from blame. He can adopt a favourable attitude toward keeping his promise and, as a result, form a desire to keep it. It's just that, were he to form such a desire, he would be unable, due of the absence of *N*, to act on this desire and decide to keep it. Thus, focusing on what Jones does in the non-actual world in which he forms a desire to keep his promise is perfectly legitimate, since the goal is to illustrate what Jones has the *option* of doing in the actual world in which he does not form this desire; and in the actual world he does have the option of continuing to deliberate in the appropriate way (and hence of forming this desire).

It would seem, then, that the alternative of continuing to deliberate cannot be ruled out by adding a time constraint: even under such a constraint, Jones has the option of deliberating in such a way that he avoids blame. Consequently, Widerker has failed, so far, to disprove PAP: a friend of the principle is entitled to ground (in part at least) Jones's responsibility for breaking his promise on this alternative. I turn now to the issue of whether the option of continuing to deliberate can be ruled out by stipulation.

2.2 The option of continuing to deliberate cannot be ruled out by stipulation

Anticipating no uncertainties or objections, Widerker simply asserts, without argument or elaboration, that the alternative of continuing to deliberate can be ruled out by stipulation. But a strong case can be made that this alternative cannot be ruled out this way: on any (reasonable) understanding of it, the stipulation either disregards Jones's libertarian freedom, which libertarian supporters of PAP will assume he possesses, undercuts his moral fitness, or rests on otherwise dubious assumptions.

Let us first construe the stipulation as the claim that Jones is an obstinate, unreasonable person, a person who, like Frankfurt's (1969, 831) Jones, 'does what he has once decided to do no matter what happens next and no matter what the cost' (cf.

for Jones to reply by saying, 'Well, I broke my promise, but I'm not blameworthy for breaking it because, in the process of deliberating about whether to keep it, it occurred to me that, if I break it, you might not bequeath to me your coin collection; and this caused me to verge on deciding to keep it.' Jones's reply runs aground, and the reason, it would seem, is that his deliberative process (regarding whether he should keep his promise) did not show the proper respect for morality.

McKenna 2003, 211). In other words, let us construe it as the stipulation that Jones is simply not the sort of agent who, in the circumstances, would continue deliberating. He knows his time is short and, given his abnormally headstrong personality, he would not let this opportunity to spend time with Mary pass him by.

This will not do, since even if Jones is unreasonable, he has the *option*, if he is libertarianly free, of continuing to deliberate (in the appropriate way). Unreasonableness about moral matters does not entail *ignorance* of moral matters. Jones, despite his assumed obstinacy, believes that he should (and can) discharge his obligation to keep his promise.

Moreover, and relatedly, it is conceptually objectionable, even assuming that his unreasonableness is psychopathological, that Jones lacks the ability to continue deliberating. Psychopathology, contrary to its conception in philosophy, does not deprive its subjects of the ability to do otherwise. As Pickard (2015), through a painstaking analysis of disorders of agency, has recently shown,¹⁹ ‘psychopathology does not strip people of free will...[A]ddicts, agoraphobics, kleptomaniacs, neurotics, obsessives, psychopathic serial murderers, and, further, patients diagnosed with disorders whose symptoms include impulsive behaviour, such as personality disorders, eating disorders, and paraphilias, have the ability to do otherwise: it is possible for them to refrain from performing the actions constitutive of the disorder’ (Pickard 2015, 137). One would have thought this was obvious; after all, people with addictions and disorders patently do refrain from performing the actions constitutive of their addiction or disorder every day, in their thousands. But a sizeable number of philosophers²⁰ seem to think that if an addiction or disorder has a sufficiently powerful grip on a person, the person is incapable of doing otherwise. This is simply false, however. Although the prognosis for addiction (for example) is, in some cases, bleak, even daily, long-term users of extremely addictive drugs like heroin, methamphetamines, and nicotine often quit.²¹ To be sure, people do sometimes lack the capacity for behavioural control, but in these cases their behaviour is ‘due to the effect on executive function of their emotional or physical state, in which case [it] approximates an automatic reflex rather than being...an action’ (*ibid*, 156). The claim being made is that there are no real-world examples of psychologically determined

19. Cf. Hyman (2007), Glannon (2008), and Heyman (2009).

20. Pickard (2015, 136) provides several examples.

21. Here is Heyman on some of the addicts featured in his book on addiction: ‘Scott was a daily methamphetamine user, then a daily heroin user; Jessie was doing cocaine at work and at home, and Patty used cocaine for fifteen years. Yet they quit’ (2009, 66).

action: we have no examples of genuine actions over which agents lack control.²² Or so the libertarian friend of PAP may plausibly argue.

But, too, if the opponent of PAP wishes to maintain that Jones is both unreasonable *and* ignorant of moral matters, or that his unreasonableness about moral matters is a consequence of his (unwilful, blameless) ignorance of moral matters, then one can reasonably rejoin that he is morally incompetent and hence not blameworthy for the wrong choices he makes by virtue this ignorance. Inculpable ignorance is exculpatory, which is why we do not (or at least should not) punish mentally disabled persons who, unbeknownst to them, act wrongly.

At any rate, Widerker himself would, I suspect, resist a reply on his behalf that appealed to Jones's being morally ignorant. Moya, in his defence of PAP against Widerker's attack on the principle, argues that Jones is not blameworthy because, due to the absence of *N*, his capacity for practical reasoning is impaired, and therefore, 'he cannot decide to keep his promise, no matter how strong or decisive the moral reasons for this decision may appear to him' (Moya 2007, 483). Widerker, in his rejoinder to Moya, maintains that even if Jones's capacity for practical decision-making is impaired, as it patently is in BMF, it does not follow that his capacity for practical reasoning is impaired. 'Moya's mistake,' he argues,

stems from thinking that an agent's sound capacity for practical reasoning entails a capacity for decision-making on the basis of reasons. An agent's capacity for practical reasoning requires that he be reasons-responsive in the sense of being able to respond differentially to reasons. However, this ability need not be cashed out in terms of *decision-making*, i.e., in terms of the different decisions the agent would be able to make (on the basis of reasons) in different circumstances. It can be cashed out in terms of his being able to have/form reasonable *beliefs* as to what he would do in various circumstances, including *beliefs* about which decisions he would make when being presented with different reasons for acting...[and] this is an ability that in [BMF] Jones *had*. (2009, 92–3)

I concur with this, and I invite the reader's concurrence with it. But note that if, on the one hand, we assume that Jones is reasons-responsive in the way described here, then we cannot, on the other hand, stipulate that he is morally ignorant, since being reasons-

22. Cf. Alvarez (2009).

responsive in this way requires being knowledgeable, at least roughly, about which reasons are pertinent (morally speaking) in the circumstances, which reasons are to be rejected, and so on.

Let us now construe the stipulation that Jones lacks the alternative of continuing to deliberate as the claim that BMF can be revised so that *N* is a causally necessary condition, not only for Jones to decide to keep his promise, but also for him to adopt a favourable attitude toward keeping it (and hence for him to form a desire to keep it). In short, let us construe it as the stipulation that *N*'s absence renders Jones incapable of deliberating further.

This will not do, either, since such a stipulation—quite apart from the fact that it concedes that BMF is not a counterexample to PAP—undermines Widerker's response to Moya's objection that, because his capacity for practical reasoning is impaired, Jones is not blameworthy. As we have seen, Widerker, in order to deflect Moya's objection, argues that, despite the fact that Jones's capacity for practical decision-making is impaired, his capacity for practical reasoning is not, since he is able to respond differentially to reasons and is thereby able to form reasonable beliefs as to what to do in various circumstances. However, if Jones cannot so much as adopt a favourable attitude toward keeping his promise, then he is not suitably reasons-responsive, since to adopt a favourable attitude toward *V* involves forming the belief that *V* is to be favoured. But Jones, by our assumption, is unable to form such a belief.

Elaborating, suppose that Jones's neurological deficit has rendered him incapable of even forming the belief that he should keep his promise. Further suppose that, while deliberating about whether to keep it, he reflects on the various (moral) reasons for keeping it (e.g., that breaking it will have bad consequences, that virtuous people keep their promises, etc.). It seems plainly false that although he is incapable of forming the belief (on the basis of these reasons) that he should keep his promise, he possesses the ability to form reasonable beliefs as to what to do in various circumstances. Trivially, the belief that he should keep his promise is reasonable, and, by our assumption, he is unable to form this belief. By Widerker's own criteria, then, of what constitutes a sound capacity for practical reasoning, Jones lacks such a capacity. But then it follows that the circumstances of the revised IRR-situation (in which *N*'s absence renders Jones incapable of even forming the belief that he should keep his promise) are such that Jones is not fit to be held responsible, since he fails to meet a requirement for responsibility of the reasons-responsiveness account of responsibility to which Widerker subscribes. Speaking more generally, to strip an agent of his ability to form reasonable beliefs as to what to do in various circumstances is effectively to strip him of his capacity to act *rationally*. But

responsibility requires this capacity, which is why a person who acts wrongly because this capacity is impaired, say because of dementia, is not blameworthy.

Finally, let us interpret the stipulation that Jones lacks the alternative of continuing to deliberate as the claim that a condition on his successfully keeping the promise is that, at the time he makes it, he must firmly commit to keeping it and agree not to subsequently deliberate about whether to keep it, say because his uncle has made this a term of their 'agreement'.²³ Call this condition *C*. If Jones agrees to *C*, then after he makes the promise, he cannot deliberate about whether to keep it without doing something morally equivalent to deciding to break it.

This interpretation of the stipulation, however, rests on an assumption that has no demand on our acceptance, namely, that it is possible to do something morally equivalent to deciding to break a promise simply by deliberating about whether to keep it. Suppose that Jones agrees to *C*, makes the promise to his uncle, and subsequently contemplates breaking it. But further suppose, counterfactually, that he has not suffered a neurological malfunction and that, in the end, he visits his uncle in the hospital as promised. Even if (contrary, as I see it, to fact) his uncle would be justified in reproaching him for contemplating breaking the promise, he would not be justified in accusing him of doing something morally equivalent to deciding to break it, since he did no such thing! Although our actions are under our voluntarily control, our motives are not: we can foster virtuous motives, but we cannot directly produce them in ourselves from one moment to the next.²⁴ And nor can we always control our thoughts, which sometimes occur without discernible impetus. One does not do something morally equivalent to deciding to break a promise if one wavers momentarily in one's commitment to it, say because one's non-virtuous motives or spontaneous thoughts have led one to contemplate breaking it.

3. An objection rebutted

In this section I confront an important objection that might be levelled against my argument that BMF does not constitute a counterexample to PAP.

3.1 The Objection

Reflection on the epistemic dimension of Jones's situation in BMF reveals that continuing to deliberate in the appropriate way is insufficiently *robust*, to use Fischer's

23. I thank an anonymous reviewer for suggesting this interpretation of the stipulation.

24. See, e.g., Ross (1930, 5).

(e.g., 1999) well-known terminology, to qualify as a morally significant alternative; it is too 'flimsy and exiguous' to ground an attribution of moral responsibility. Jones is cognitively insensitive to the fact that continuing to deliberate in the appropriate way exempts him from blame. Widerker has adduced an example, call it POTION, in which near Jones is a cup of water that, unbeknownst to him, contains a sleeping potion. It is up to Jones whether he drinks the water, but this alternative is irrelevant to his moral responsibility, since he is cognitively insensitive to the fact that by drinking the water he can avoid breaking his promise.²⁵ 'In such a situation,' Widerker avers, 'it would be counterintuitive to ground Jones's culpability for D(-K) in the fact that he did not avail himself of the alternative possibility of drinking the said cup of water' (2006, 175).²⁶ Similarly, says our objector, it would be counterintuitive to ground Jones's culpability for D(-K) in the fact that he did not avail himself of the option of continuing to deliberate. Succinctly put, the option of continuing to deliberate fails to meet the requirements for a morally significant alternative. As I have noted, Widerker defines a morally significant alternative as follows:

(MSA₁): [A]n act V* [is] a morally significant alternative to a morally wrong act V that an agent S performs at t if and only if: (i) V and V* are incompatible, (ii) S is aware (or should be aware) that if he performs V* at t, he will not be acting in a morally wrong way, (iii) S truly believes that it is within his power to perform V* at t, and (iv) in the circumstances, V*-ing at t would be regarded as a reasonable way for S to avoid V-ing at t. (2006, 175)²⁷

Observe that the option of continuing to deliberate fails to meet the second (and possibly the fourth) of these requirements.

Reply: We may dispute the assumption that Widerker's definition of a morally significant alternative (MSA₁) is adequate. Our objector assumes that it is, and then

25. See also Pereboom (2003, esp. 187–8).

26. Strictly speaking, drinking the water is a morally significant alternative, since breaking the promise has moral value, and if V has moral value, then any action with a different moral value will be morally better or morally worse than V and thereby be significant. As McKenna has pointed out, however, 'within the very wide spectrum of courses of action with differing moral weights or values, only some in a deliberative context will be relevant to *competent moral deliberation and agency*' (2003, 207). Drinking the water is not, in this latter sense, a morally significant alternative.

27. I have altered Widerker's notation slightly.

claims, in essence, that any alternative that fails to satisfy its requirements has to be jettisoned. But why should we accept this line of reasoning? Why should we not assume, instead, that NB is true and that continuing to deliberate is actionally accessible to Jones, and then claim that since MSA_1 excludes this alternative, MSA_1 has to be revised? After all, MSA_1 is based solely on its purported intuitive resonance,²⁸ and since the alternative of continuing to deliberate seems intuitively to be morally significant, it is reasonable to conclude that MSA_1 requires revision.

These reflections suggest that we need to construct a definition of a morally significant alternative that, while it excludes from its extension alternatives like the one described in POTION, includes in its extension alternatives like the one described in BMF*. This can be accomplished by augmenting MSA_1 as follows:

If, through no fault of her own, S is prevented from doing V^* despite doing everything in her power to do V^* , then S 's doing everything in her power to do V^* is a morally significant alternative to doing V .

As the reader will notice, this is a variation on NB. If this clause is appended to MSA_1 , continuing to deliberate in the appropriate way is a morally significant alternative (where continuing to deliberate constitutes Jones's doing everything in his power to decide to keep his promise), whereas alternatives like the one described in POTION are not.

But there is another reason to reject MSA_1 that is independent of the fact that it conflicts with my contention that the option of continuing to deliberate is a morally significant alternative. As previously discussed, not all wrong actions are equally bad (just as not all right actions are equally good), so the degree of blameworthiness we attribute to an agent who has performed a wrong action will depend on the wrongfulness of the action. While Jones, then, is perhaps blameworthy for momentarily contemplating breaking his promise in the first place,²⁹ he is not *as* blameworthy as he would have been had he decided outright to break it. Thus, he could have acted appreciably less badly than he did (while still doing something blameworthy), something he knew (we may assume). But MSA_1 cannot capture this intuitive fact, for if MSA_1 is correct, then an

28. It is also noteworthy that there are competing definitions in the literature. McKenna (2003, 209) and Pereboom (2003, 188), for example, have offered definitions that Widerker finds unsatisfactory.

29. Though, again, we cannot rule out the possibility that, while he was deliberating, Jones never so much as entertains the idea of breaking his promise, that his whole deliberative process (prior to t) consists of his repeatedly verging on deciding outright to keep it (only to be foiled by the absence of N). Since we cannot rule out this possibility, Jones has an alternative that allows him to avoid blame altogether.

agent has to be aware (or should be aware) that if he performs V^* (a morally significant alternative to a wrong act V), 'he will not be acting in a morally wrong way'. With this in mind, consider the following slightly modified version of MSA_1 :

(MSA_2): [A]n act V^* [is] a morally significant alternative to a morally wrong act V that an agent S performs at t if and only if: (i) V and V^* are incompatible, (ii) S is aware (or should be aware) that if he performs V^* at t , he will be acting less badly than if he does V , (iii) S truly believes that it is within his power to perform V^* at t , and (iv) in the circumstances, V^* -ing at t would be regarded as a reasonable way for S to avoid V -ing at t .

This version of a morally significant alternative retains what is attractive about MSA_1 , and it includes in its extension the alternative of continuing to deliberate in the appropriate way (while excluding alternatives like the one described in POTION), but it is superior to MSA_1 because, unlike MSA_1 , it accommodates the intuition that a morally significant alternative to a wrong action need not be an action that the agent believes, or should believe, is itself not morally wrong.

In any case, the strategies I have adumbrated—viz., augmenting MSA_1 or replacing it with MSA_2 —are reasonable, which means that the PAP enthusiast need not acquiesce to the claim that since the alternative of continuing to deliberate fails to meet the requirements of a morally significant alternative as specified in MSA_1 , the alternative of continuing to deliberate is not morally significant.

4. Conclusion

I conclude that BMF is not an IRR-situation and, consequently, that Widerker has failed to generate a convincing argument against PAP. Attempts to overturn PAP by dint of counterexample have not fared well—certainly there is nothing resembling a consensus regarding whether the myriad fanciful counterexamples on offer succeed—and, as it turns out, Widerker's proposed counterexample, as ingenious and initially compelling as it is, is just another putative counterexample in a long line of putative counterexamples that, on closer inspection, fails to demonstrate that the principle is false.

References

- Alvarez, M. 2009. "Actions, Thought-Experiments and the 'Principle of Alternate Possibilities.'" *Australasian Journal of Philosophy* 87: 61–81.
- Fischer, J. M. 1999. "Recent Work on Moral Responsibility." *Ethics* 110: 93–139.
- Fischer, J. M. 2010. "The Frankfurt Cases: The Moral of the Story." *Philosophical Review* 119: 315–36.
- Frankfurt, H. 1969. "Alternate Possibilities and Moral Responsibility." *Journal of Philosophy* 66: 829–39.
- Ginet, C. 1990. *On Action*. New York: Cambridge University Press.
- 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.
- Haji, I. and McKenna, M. 2004. "Dialectical Delicacies in the Debate about Freedom and Alternative Possibilities." *Journal of Philosophy* 101: 299–314.
- Glannon, W. 2008. "Moral Responsibility and the Psychopath." *Neuroethics* 1: 158–66.
- Heyman, G. M. 2009. *Addiction: A Disorder of Choice*. Harvard: Harvard University Press.
- Hyman, S. E. 2007. "The Neurobiology of Addiction: Implications for Voluntary Control of Behavior." *The American Journal of Bioethics* 7: 8–11.
- Kane, R. 1996. *The Significance of Free Will*. New York: Oxford University Press.
- Larvor, B. 2010. "Frankfurt Counter-Example Defused." *Analysis* 70: 506–8.
- McKenna, M. 2003. "Robustness, Control, and the Demand for Morally Significant Alternatives: Frankfurt Examples with Oodles and Oodles of Alternatives." In *Moral Responsibility and Alternative Possibilities: Essays on the Importance of Alternative Possibilities*, edited by Widerker and McKenna, 201–17. Aldershot, UK: Ashgate.
- Mele, A. and Robb, D. 1998. "Rescuing Frankfurt-Style Cases." *Philosophical Review* 107: 97–112.
- Moya, C. J. 2007. "Moral Responsibility without Alternative Possibilities?" *Journal of Philosophy* 104: 475–86.
- Pereboom, D. 2003. "Source Incompatibilism and Alternative Possibilities." In *Moral Responsibility and Alternative Possibilities: Essays on the Importance of Alternative Possibilities*, edited by Widerker and McKenna, 185–99. Aldershot, UK: Ashgate.

- Pickard, H. 2015. "Psychopathology and the Ability to Do Otherwise." *Philosophy and Phenomenological Research* 90: 135–63.
- Ross, W. D. 1930. *The Right and the Good*. Oxford: Clarendon Press.
- van Inwagen, P. 1983. *An Essay on Free Will*. Oxford: Clarendon Press.
- Widerker, D. 1991. "Frankfurt on 'Ought Implies Can' and Alternative Possibilities." *Analysis* 51: 222–4.
- 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 the Principle of 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 *Moral Responsibility and Alternative Possibilities: Essays on the Importance of Alternative Possibilities*, edited by Widerker and McKenna, 53–73. Aldershot, UK: Ashgate.
- Widerker, D. 2005. "Blameworthiness, Non-robust Alternatives, and the Principle of Alternative Expectations." *Midwest Studies in Philosophy* 29: 292–306.
- Widerker, D. 2006. "Libertarianism and the Philosophical Significance of Frankfurt Scenarios." *Journal of Philosophy* 103: 169–87.
- Widerker, D. 2009. "A Defense of Frankfurt-friendly Libertarianism." *Philosophical Explorations* 12: 87–108.
- Widerker, D. and Goetz, S. 2013. "Fischer against the Dilemma Defence: The Defence Prevails." *Analysis* 73: 283–95.
- Widerker, D. and McKenna, M. (eds.). 2003. *Moral Responsibility and Alternative Possibilities: Essays on the Importance of Alternative Possibilities*. Aldershot, UK: Ashgate.