

Frankfurt cases and the (in)significance of timing: a defense of the buffering strategy

Author(s): David Hunt and Seth Shabo

Source: *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, July 2013, Vol. 164, No. 3 (July 2013), pp. 599-622

Published by: Springer

Stable URL: <https://www.jstor.org/stable/41932749>

---

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



JSTOR

Springer is collaborating with JSTOR to digitize, preserve and extend access to *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*

## Frankfurt cases and the (in)significance of timing: a defense of the buffering strategy

David Hunt · Seth Shabo

Published online: 2 March 2012  
© Springer Science+Business Media B.V. 2012

**Abstract** Frankfurt cases are purported counterexamples to the Principle of Alternative Possibilities, which implies that we are not morally responsible for unavoidable actions. A major permutation of the counterexample strategy features buffered alternatives; this permutation is designed to overcome an influential defense of the Principle of Alternative Possibilities. Here we defend the buffering strategy against two recent objections, both of which stress the timing of an agent’s decision. We argue that attributions of moral responsibility aren’t time-sensitive in the way the objectors suppose. We then turn to the crucial question of when an action is relevantly avoidable—when, in the parlance of the literature, an alternative possibility is robust. We call attention to two plausible tests for robustness that merit further consideration, showing that the agents in buffered Frankfurt cases don’t pass these tests, despite being morally responsible for their actions.

**Keywords** Frankfurt cases · Buffered alternatives · Franklin · Palmer · Hunt · Pereboom

Some 40 years ago, Harry Frankfurt presented an apparent counterexample to the principle of alternative possibilities (PAP), which states that “someone is morally responsible for what he has done only if he could have done otherwise” (Frankfurt 1969, p. 829). To disprove this principle, Frankfurt described a scenario in which the circumstances that render a particular action unavoidable play no role in actually

---

D. Hunt  
Philosophy Department, Whittier College, Whittier, CA 90608, USA  
e-mail: dhunt@whittier.edu

S. Shabo (✉)  
Philosophy Department, University of Delaware, Newark, DE 19716, USA  
e-mail: sshabo@udel.edu

 Springer

producing that action. Since the agent in Frankfurt's scenario acts entirely on his own, as if nothing prevented him from doing otherwise, it's natural to conclude with Frankfurt that he is morally responsible for that action, unavoidable though it is. If so, PAP is false.

Frankfurt's attack on PAP has sparked a sizable literature, in which friends of the counterexample strategy have sought to overcome a number of subtle and thoughtful defenses of PAP. Here we examine one influential extension of Frankfurt's challenge, along with the defense of PAP to which it is addressed. The extension of Frankfurt's challenge features "buffered" scenarios, Frankfurt cases in which the agent must complete an intermediate step (traverse a psychological buffer, so to speak) before she is in a position to do otherwise—specifically, before she can *decide* upon an alternative course of action. The buffering strategy, as we shall call it, is addressed to PAP-defenders who deny that someone can be morally responsible for an unavoidable *decision*, given a plausible understanding of decisions as mental acts, together with an incompatibilist account of responsible agency. This response to Frankfurt has come to be known as 'the Kane–Widerker–Ginet defense' (Kane 1985, p. 51; Kane 1996, pp. 142–143, pp. 191–192; Widerker 1995; Ginet 1996); if it succeeds, incompatibilist defenders of PAP will have met Frankfurt's challenge. Against the Kane–Widerker–Ginet defense, buffering strategists claim that agents can be morally responsible for unavoidable decisions, even granting incompatibilism and their opponents' view of decisions, provided that the decision to do otherwise is a buffered alternative.

We believe that the buffering strategy succeeds against the Kane–Widerker–Ginet defense. Our aim is to show that two recent objections to the buffering strategy fail. Both of these objections (Palmer 2011; Franklin 2011) make much of the exact time to which attributions of moral responsibility pertain. We argue that ordinary attributions of moral responsibility are not time-sensitive in the way these objections suppose. Indeed, we believe that there is very little to be said for this idea. Nevertheless, the fact that two recent defenders of PAP have independently revived this idea (a version of which was first proposed by Carl Ginet<sup>1</sup>) shows that buffering strategists must take it seriously. If we are right, the "exact timing defense" poses no real threat to the buffering strategy.

In Sect. 1, we review the state of play between the PAP-defender and the buffering strategist. In Sect. 2, we introduce the exact timing defense, and in Sects. 3 and 4 we take up Palmer's and Franklin's respective versions of the defense. As will be seen, both authors argue that the agent in buffered Frankfurt cases is morally responsible only for deciding as he does at a precise time, and that this outcome is relevantly avoidable; hence no single item is both unavoidable and something for which the agent is morally responsible, contrary to what buffering strategists maintain. In Sect. 5, we explore the much-discussed question of when an action is relevantly avoidable—when, in the parlance of the literature, the agent has a "robust" alternative possibility to performing that action (see especially Fischer

<sup>1</sup> Ginet (2002, pp. 307–308) briefly sets out a version of the timing objection against Pereboom's (2001) buffering strategy. This objection is an adaptation of one of Ginet's (1996, pp. 406–409) earlier responses to Frankfurt's original counterexample. Because it is very brief, it is hard to say exactly how much Ginet's version of the timing objection has in common with Palmer's and Franklin's respective versions.

1994, chap. 7). We propose two answers that are different from the one that Palmer and Franklin employ in their critiques of the buffering strategy.

## 1 Frankfurt cases and the buffering strategy

In Frankfurt's original counterexample, Jones decides to perform an action (kill Smith, say), and follows through on this decision, under conditions that would ordinarily make him morally responsible for his action, if anyone is ever morally responsible for an action. But this isn't an ordinary case: lurking in the background is a third party, Black, equipped with a sci-fi device programmed in such a way that the following are both true:

- (1) If Jones were going to decide to kill Smith, the device would leave Jones alone.
- (2) If Jones were *not* (otherwise) going to decide to kill Smith, the device would intervene to bring it about that Jones *does* decide to kill Smith.

In Frankfurt's example it's the antecedent of (1) that happens to be true, so the device leaves Jones alone and he decides on his own. It's (1) that's supposed to preserve Jones's prima facie moral responsibility for deciding to kill Smith, while (2) is supposed to ensure that Jones can't do otherwise than decide to kill Smith.

The coherence of this case clearly requires that Black and his device have some way of telling whether it's the antecedent of (1) or of (2) that is satisfied; otherwise the device won't be able to function as advertised. Frankfurt had simply noted (parenthetically) that "Black is an excellent judge of such things" (op. cit., p. 835). But however excellent Black's judgment, it must be based on some distinguishing feature of the case. If Jones were going to decide to kill Smith, there must be some sign S1 indicating this, so that on S1's appearance Black can be confident that intervention is unnecessary; likewise, if Jones were *not* going to decide to kill Smith, there must be some sign S2 indicating this, so that on S2's appearance the device can intervene, ensuring that Jones decides to kill Smith after all. Prior signs became a prominent feature of the literature soon after the appearance of Frankfurt's original article (beginning with Blumenfeld 1971).

Once the centrality of prior signs was recognized, however, a powerful argument became available to PAP's defenders.<sup>2</sup> Frankfurt's counterexample, in which an agent is supposed to be morally responsible for an action despite being unable to do otherwise, includes the appearance of S1, the consequent nonintervention by Black's device, and Jones's performing the action Black wants him to perform. But the relation between S1 and Jones's action gives rise to a dilemma. Either S1 (or the underlying conditions it manifests) causally determines Jones's action, or it doesn't. If it does, then incompatibilists will deny that Jones's action was free in the sense required for moral responsibility. But if S1 *doesn't* causally determine Jones's action, there is no reason to think that Jones can't do otherwise. (Black's device

<sup>2</sup> David Widerker and Michael McKenna call this "the most compelling of the defenses of PAP" (2003, p. 8).

would intervene to force the action in question on Jones if and only if S2 appeared; what actually appeared was S1, and S1, on the second horn of the dilemma, leaves it causally open that Jones might do otherwise.) In short, Frankfurt cases either illicitly assume that someone can be morally responsible for a causally determined action, or they fail to show that the agent can't do otherwise.

The Kane–Widerker–Ginet defense is the most influential version of this response. These writers make it explicit that *decisions*, not outward actions, are the fundamental *loci* of moral responsibility (see especially Widerker 1995, p. 247). In their view, Frankfurt cases fail to show that someone can be both morally responsible for, and unable to avoid, her causally undetermined decision. Thus a version of PAP survives the counterexample strategy; according to the modified principle, someone is morally responsible for deciding as he has only if he could have decided otherwise.

The Kane–Widerker–Ginet defense (henceforth ‘the dilemma defense’) can be challenged from multiple directions. One challenge involves buffered alternatives. Before coming to this, it will be helpful to examine a related response to the dilemma defense. This response relies on the view that decisions have a *complex* temporal structure; that is, decisions have an initial “part” or stage, and this part can serve as the sign that tells Black whether or not to intervene (Stump 1996, 1999). This would be an *internal* rather than a *prior* sign of the decision, and it would allow the critic of PAP to pass between the horns of the dilemma defense. The PAP-defender cannot object that the prior sign causally determines Jones’s decision, for there is no *prior* sign. Nor can the PAP-defender maintain that Jones is able to do otherwise, given that his decision is causally undetermined; for any alternative decision would be announced by the appearance of its initial part, thereby triggering Black’s intervention. In sum, Jones’s actual decision can be causally undetermined, as required by incompatibilists, while remaining unavoidable, inasmuch as Black’s device would nip any alternative decision in the bud.

Among proponents of the dilemma defense, Widerker has been particularly alert to the threat posed by the foregoing challenge, to which he has offered two responses. First, he notes that the defender of PAP can simply shift his focus to the initial part of the decision. The agent remains free to *initiate*, if not complete, an alternative decision, and this is arguably enough for the agent to satisfy PAP’s condition for moral responsibility (1995, pp. 252–253). Second, he denies the complexity premise itself: “a decision, being the forming of an intention, is a *simple* mental action that does not exhibit the sort of complex structure assumed by the objector” (p. 253). While we might say that Jones is in the process of *trying* to decide whether to kill Smith, we would not say that he is in the process of *deciding* to kill Smith. If this is right, it appears that Jones *is* able to do otherwise; that is, he is able to *begin* to make a different, causally undetermined decision before Black can intervene, and his decision, being internally simple, ends when it begins. Thus, Widerker concludes, a version of PAP restricted to acts of deciding survives Frankfurt’s challenge.<sup>3</sup>

<sup>3</sup> According to the restricted principle, PAPD, “A person is morally responsible for his *decision* (choice, undertaking) to do A only if he could have decided otherwise” (p. 259, emphasis original).

We can now introduce the buffering strategy. This strategy grants, at least for the sake of argument, both of Widerker's contentions in the preceding paragraph. That is, it grants both (i) that the initial part of a decision, if decisions did have parts, might be sufficiently robust to satisfy PAP, and (ii) that decisions are best thought of as simple, not complex. It therefore adopts a *prior* rather than an *internal* sign approach. But its prior sign also functions in some ways like the initial part of the decision under the internal sign approach. On that approach, for an alternative decision to occur, its initial part must first occur; the initial part is a *necessary condition* for the decision. Alternatively, we might say that it's the initial part of the alternative decision to which the agent has *immediate access*; he has access to the alternative decision only by first accessing its initial part. The same is true of the buffers in buffer cases. These buffering events or states are distinct from the decision; they occur or obtain prior to the decision. But they are a necessary condition for the decision, which cannot be accessed except via the buffer. Entering the "buffer zone" would trigger the counterfactual device, making it impossible for the agent to access an alternative decision. But in buffering counterexamples the agent doesn't actually enter the buffer zone; consequently, as in every Frankfurt scenario, the device remains quiescent, the agent decides on his own, and our judgment is that he is morally responsible for his decision.

The originators of the buffering strategy are David Hunt and Derk Pereboom.<sup>4</sup> Hunt (2005, pp. 132–34) deploys a variation on Frankfurt's original story that he calls *Revenge*. Jones finds himself entertaining the idea of killing Smith, who has just insulted him. The idea grows in strength, culminating in Jones's decision to kill Smith. The conditions under which Jones acts (leaving the counterfactual intervener out of the account for the moment) are such that he could have decided otherwise, in the incompatibilist's sense of 'could'; but he would first have needed to entertain or consider this alternative, and he doesn't in fact do so. Had he done so, Black would have intervened and brought about the decision to kill Smith. Here Jones's considering not killing Smith serves as a buffer zone between his actual course of (mental) action and the morally relevant alternative of deciding not to kill Smith. Since entering the buffer zone would trigger the intervener, Jones cannot access a morally relevant alternative, yet he seems morally responsible for his actual decision, which he reaches on his own.

In Pereboom's *Tax Evasion 2* (loc. cit.), Joe decides to cheat on his taxes and couldn't have decided otherwise without first achieving a certain level of attentiveness to his moral reasons. This attentiveness is causally necessary but *not* causally sufficient for his deciding not to cheat on his taxes. Were Joe to achieve this attentiveness, Black would hijack his mental processes and ensure that he decide to cheat on his taxes. As in *Revenge*, we have a buffer zone separating the agent from a morally relevant alternative; had the agent entered the buffer zone, he would have

---

<sup>4</sup> Hunt and Pereboom came up with the idea of buffered alternatives independently. Hunt proposed a buffer case in Hunt (2000, p. 215), but developed it into a full-fledged argument against PAP only in Hunt (2005). Pereboom has revised his original example, *Tax Evasion* (2000, p. 30, 2001, pp. 18–22); we focus here on *Tax Evasion 2* (2005, pp. 231–232). For an application of the buffering strategy to Kane's (2000) "dual regress" defense of an avoidability requirement, see Shabo (2010, pp. 362–365).

been forced back on track, but because he doesn't actually enter the zone his decision is his own and it seems that he is morally responsible for it.

Essential to both Hunt's and Pereboom's cases is the stipulation of a mental act (entertaining a course of action, attending to a moral reason) that does not itself constitute the kind of alternative called for under PAP, but rather serves as a necessary condition for access to such an alternative. Hunt and Pereboom needn't claim that every action or decision has buffers of this sort; a single counterexample is enough to refute PAP, and all that's required is that the buffers in *Revenge* and *Tax Evasion 2* be psychologically plausible and that positing them doesn't undermine the agents' responsibility for what they actually decide.<sup>5</sup> If this is right, the buffering strategy can provide counterexamples to PAP that don't fall prey to the dilemma defense.

## 2 The exact timing of decisions

As noted, the buffering strategist grants both that decisions are simple mental acts, and that the agent's actual decision in *Revenge* and *Tax Evasion 2* isn't causally determined by his antecedent mental states. A key element of the buffering strategy is the claim that decisions (simple or otherwise) need not be *immediately accessible* to an agent. At the precise moment when someone decides as he actually does, the idea is, it may not be open to him to decide differently, but only to *prepare* himself to decide differently in the next instant; the alternative decision becomes accessible only when this preliminary step is completed. Of course, given the presence of the counterfactual intervener, taking this step won't actually make the alternative decision accessible. Since the dilemma defense equates having a robust alternative with being able to decide to do otherwise, buffered Frankfurt cases appear to circumvent this defense. The agent seems to be morally responsible for the decision he makes, despite lacking a robust alternative to it, simple and causally undetermined though it is.

Against the buffering strategy, Franklin and Palmer argue that a robust alternative possibility remains in these cases, owing to the specific timing of the events involved. Recall *Tax Evasion 2*. Since Joe's actual decision to evade taxes was causally undetermined, it seems that Black was unable to know in advance that Joe would make it on his own. Black could ensure that (one way or another) Joe would decide to evade taxes because Black knew that Joe must enter the buffer zone (that is, pay more attention to his moral reasons) before deciding otherwise, and because Joe's entering this zone would have triggered Black's intervention. What Black could *not* have ensured, according to Franklin and Palmer, is that Joe would decide to evade taxes exactly when he did. For it was within Joe's power to have entered the buffer zone at the very moment when he actually decided to evade taxes. And, the thought continues, since Black's resulting intervention would take time, the Black-induced decision would have to occur *somewhat later* than Joe's actual

<sup>5</sup> Since PAP is a conceptual thesis, psychological plausibility is not strictly required, though examples involving recognizably human agents will undoubtedly have more persuasive force.



decision. Whatever else, then, Joe could have avoided deciding to evade taxes when he actually did so decide; and, Franklin and Palmer maintain, this alternative possibility satisfies Pereboom's requirement for robustness (cf. Ginet 2002, pp. 307–308). If well founded, this response generalizes to other buffered Frankfurt cases.

We shall come to Pereboom's requirement in due course. First, however, it should be stressed that identifying a plausible candidate for a robust alternative possibility (that is, one that seems capable of grounding the agent's responsibility in *some* Frankfurt case or other) is only half the battle for opponents of the buffering strategy. To see this, suppose that Franklin and Palmer were to say that Joe is morally responsible for *deciding to evade taxes*. Even if we grant for argument's sake that deciding to evade taxes *at a later time* is a plausible candidate for a robust alternative possibility, this won't be enough to save PAP. For, contrary to what PAP implies, Joe will *actually* be morally responsible for something he is unable to avoid doing (and that he is unable to avoid responsibility for doing): deciding to evade taxes. What is needed, then, is a robust alternative possibility that "fits" what the agent is actually responsible for; along with being capable of grounding an agent's responsibility, such an alternative possibility will allow the PAP-defender to maintain that the agent is able to avoid the very thing for which he is morally responsible (or at least his *being* morally responsible for that thing).

Palmer and Franklin both maintain that the agents in buffered Frankfurt cases have a robust alternative possibility in virtue of being able to avoid deciding exactly when they do. For reasons indicated in the previous paragraph, this line of response won't work if the agents in these cases are morally responsible for deciding as they do *simpliciter* (as opposed to so deciding at a precise time). It is not surprising, then, that Palmer and Franklin seek to restrict *what* agents are morally responsible for in buffered Frankfurt cases. In their view, such agents aren't morally responsible for deciding as they do *simpliciter*,<sup>6</sup> but only for so deciding *exactly when they do*. In *Tax Evasion 2*, if the alternative sequence (in which Joe's decision comes at a later time) constitutes a robust alternative possibility, PAP will be safe: there will be no single item such that Joe is both morally responsible for and unable to avoid (responsibility for) it.

Defenders of the buffering strategy have two main avenues of response available. One is to challenge the restriction on what the agent is morally responsible for. If, contra Palmer and Franklin, Joe is morally responsible for deciding to evade taxes *simpliciter*, and not just for so deciding at a precise time, this line of objection fails. The second approach is to dispute the robustness of the alleged alternative possibility. If Joe's alternative possibility doesn't satisfy Pereboom's requirement for robustness, suitably understood, the objection fails for a different reason. We shall address the first response to Palmer's version of the objection and the second to Franklin's. As will become clear, however, we believe that each response is ultimately effective against both versions.

---

<sup>6</sup> Palmer is more explicit about the restriction than Franklin is; as will be seen, however, both are committed to some version of it.



### 3 Palmer's objection to Pereboom

Against Palmer's temporal restriction on Joe's moral responsibility, we shall argue that such a restriction is tenable in Frankfurt cases only if it's tenable outside of Frankfurt cases, and that it isn't tenable outside of Frankfurt cases.

Consider, then, Palmer's claim that what Joe is morally responsible for in *Tax Evasion 2* is not deciding to evade taxes, but only for so deciding when he actually does. Palmer is not explicit about whether he believes that agents are *never* morally responsible for what they do (but only for doing it when they do) or whether he takes the restriction to be a *special* feature of unavoidable action. He writes:

The key point of difference between [opponents and defenders of the buffering strategy] must center on whether or not the fact that Joe made the decision at that time is an *essential* or *necessary* part of a correct description of the thing for which he is morally responsible...Pereboom does nothing to substantiate his unease with this assumption (other than to register his opposition to it), so we are left with no answer from him as to why it might be wrong to index responsibility to precise times, why it might be wrong to think that our moral responsibility for our actions is necessarily tied to the very times at which we perform those actions (2011, p. 268).

Here Palmer seems to suppose that it's incumbent on Pereboom to explain why we are morally responsible for our actions *simpliciter*, and not just for performing them exactly when we do. We believe that this supposition is mistaken.<sup>7</sup> To see why, consider an ordinary case of *avoidable* action:

(Z1) Zeke has broken into Chad's office, where he hopes to surprise and kill Chad as part of his plan to eliminate a key witness in the case against him. When Chad enters the office alone a short while later, Zeke steps out from the shadows, gun in hand. Suppressing pangs of conscience, he stares into the other man's eyes for a full second before deciding to pull the trigger without further ado.

In Z1, Zeke's outward act of pulling the trigger proceeds in a normal way from his decision to pull it then, and this decision is causally undetermined by his proximal mental states. Furthermore, up until this decision occurs, Zeke is free to do otherwise. He is able, for example, to relent and settle for witness intimidation instead of elimination, or simply to postpone deciding for a short but indefinite period.

Imagine now that Zeke realizes a permutation of this last possibility:

(Z2) Everything is the same as in Z1, only Zeke stares into Chad's eyes for a full *two* seconds before deciding to pull the trigger without further delay.

<sup>7</sup> Indeed, we believe that it would be incumbent on Palmer to justify even the weaker supposition that we are typically morally responsible only for deciding as we do *during very brief intervals*, let alone at utterly specific instants.

In Z2, the decision and outward action occur exactly 1 second later than do their counterparts in Z1. It seems clear not only that Zeke is equally blameworthy for his conduct in the two scenarios, but that he is morally responsible (and blameworthy) for the same important thing: deciding to shoot Chad. Whether Zeke so decides at  $t$  or at  $t + 1$  (1 second after  $t$ ) doesn't change this.

This is not to deny that Zeke is morally responsible for so deciding at  $t$  in Z1 and at  $t + 1$  in Z2. Our contention is that these further, temporally specific attributions of moral responsibility are of little account. Barring special assumptions, once it's been noted that Zeke is morally responsible for deciding to shoot Chad as part of his plan to eliminate a witness, noting that Zeke is *also* morally responsible for so deciding at  $t$  in Z1 (and for so deciding at  $t + 1$  in Z2) adds little if anything of consequence.

To further underscore that the exact time doesn't matter, we might imagine that in Z3, the situation plays out just as it does in Z1, except that Chad, having decided not to wait for the slow elevator in his building, arrives in his office 10 seconds earlier, with the result that Zeke decides at  $t - 10$  (10 seconds before  $t$ ) to shoot Chad straightaway. Once again, the important thing for which Zeke is morally responsible is deciding to shoot Chad in these circumstances, not for so deciding at  $t - 10$ .

We have argued that, where *avoidable* actions are concerned, agents are typically morally responsible not only for making decisions at the precise times they do, but also—and much more importantly—for deciding as they do in the circumstances they are in, regardless of exact timing. If Palmer wishes to deny that Zeke is morally responsible for deciding as he does *simpliciter* in Z1–Z3, it's up to him to say why. Alternatively, Palmer might allow that Zeke is morally responsible for so deciding *simpliciter*, and that the temporal restriction applies only in cases of unavoidable action.

To repeat, Palmer is not explicit about whether he takes the temporal restriction to be a *special* feature of unavoidable action. Be this as it may, the only reason for the restriction he offers is particular to such action:

In my view, while Joe can be responsible for deciding to evade taxes at  $t1$ , he *cannot* be responsible for the more general fact that he decided to evade taxes *simpliciter*. This is because, due to the neuroscientist's presence, this more general fact is not something that Joe could have avoided (p. 269).

What absolves Joe of moral responsibility for deciding to evade taxes, Palmer suggests, is precisely that the neuroscientist's presence renders this outcome unavoidable. If this is indeed the *sole* reason for restricting the content of Joe's responsibility, Joe and Zeke will be in different situations. Whereas Zeke will be morally responsible for deciding as he does, Joe will be morally responsible only for what he can avoid: deciding as he does at  $t1$ .

Now there is nothing to prevent a committed defender of PAP from insisting that avoidability constrains moral responsibility in precisely this way. But why believe it? After all, Frankfurt cases provide apparent counterexamples to the view that we are morally responsible only for what we can avoid.<sup>8</sup> In *Tax Evasion 2*, Joe certainly

<sup>8</sup> As Pereboom (forthcoming) puts it, responses to such apparent counterexamples are dialectically unsatisfying if they must appeal to the very intuitions that the counterexamples are meant to challenge, rather than providing independent grounds for doubting the counterexamples.

*seems* to be morally responsible for deciding to evade taxes,<sup>9</sup> notwithstanding his inability to avoid this outcome. Supposing that Joe *would* be morally responsible for so deciding in Black's absence—something that Z1–Z3 strongly suggest, and that Palmer doesn't deny—it's hard to see why Black's presence should change this. As with Zeke in Z1–Z3, it seems that the important thing for which Joe is morally responsible is deciding as he does, not for so deciding exactly when he does. Apart from the fact that restricting the content of Joe's responsibility in *Tax Evasion 2* provides a way out for defenders of PAP, we see nothing to recommend this restriction. It's one thing to observe that the restriction helps to accommodate apparent counterexamples, another to show that the restriction is well motivated.

The *prima facie* moral triviality of the timing of Zeke's decision is of a piece with the moral triviality of other properties of Zeke's decision, making it all the more incumbent on Palmer to explain why the exact time at which an action is performed should play the unique role he assigns to it. In Z1, for example, Zeke not only decides *at t* to pull the trigger; he also so decides *from behind the desk, standing ten feet from Chad, with a bead of sweat running down the bridge of his nose, as a car backfires on the street*, while in Z2 he not only decides *at t + 1* to pull the trigger; he also decides *near the filing cabinet, standing nine feet from Chad, with a bead of sweat hanging from the tip of his nose, as a second backfire is heard from the street*. Just as Zeke is morally responsible in Z1 for deciding at a certain point in time (*t*) to shoot Chad, so he is morally responsible for deciding at a certain point in space (behind the desk, ten feet from Chad, and so on) to shoot him. But it would be absurd to suppose that in ordinary cases of avoidable action the *location* of the decision is an essential part of what the agent is morally responsible for, and that in a Frankfurt version of Z1 Zeke wouldn't be responsible for deciding to shoot Chad *simpliciter* but only for so deciding at the exact location where the decision is made. Given that this assignment of responsibility lacks any independent plausibility, the fact that it would blunt the buffer strategist's attack on PAP simply underscores the *ad hoc* nature of this move. We believe that the onus is on Palmer to explain why the parallel move, privileging the exact time of an action, is not similarly *ad hoc*. If there is a plausible way to explain this, we don't see it.<sup>10</sup>

We have argued that the precise instant at which a decision occurs is typically inessential to what the deciding agent is morally responsible for: an agent who is morally responsible for deciding to A at *t* is also, in normal cases (we'll look at

<sup>9</sup> Moreover, along with being morally responsible for deciding to evade taxes, Joe seems to be morally responsible (*pace* Palmer) for so deciding during some longer time period—say, the interval during which he had the opportunity to make and directly implement the decision (Cf. Pereboom 2001, p. 30).

<sup>10</sup> One thought that has been suggested to us is that moral responsibility for an action is temporally specific because the exact time at which an act-token occurs is essential to that act-token. A discussion of this view and why we find it implausible is beyond our present scope. Suffice it to say that a version of the timing objection that proceeded from this premise would look very different from the one Palmer actually offers. As to why Joe is morally responsible only for his decision at *t* to evade taxes, rather than for so deciding *simpliciter*, such a version would imply that this is because *t* is the only time at which that decision-token *could* have occurred. Nothing in Palmer's discussion clearly suggests that this is what he had in mind. Moreover, if exact timing *were* essential to act-tokens, we would have reason to believe, in light of cases like Z1–Z3, that act-tokens (including decision-tokens) are *not* what we are mainly interested in when we assign moral responsibility.

some abnormal cases in a moment), morally responsible for deciding to A *simpliciter*; and we see no good reason to think that introducing a Frankfurt intervener would make any difference to this. But not only is timing typically trivial, morally speaking; in some cases it isn't an object of moral responsibility at all. In Z4, Chad's building is heavily guarded and Zeke waits outside, hoping for an opportunity to get a clear shot at Chad, should he pass by the one window whose blinds happen to be open. His sniper's rifle trained on that window, Zeke gets lucky: Chad appears momentarily in the window and Zeke immediately squeezes off a shot, killing Chad instantly. Zeke is morally responsible for many things here: for accepting the contract to kill Chad; for lying in wait with intent to kill; for killing Chad; and also for deciding, at the exact time Chad appeared in the window, to fire a gun aimed at Chad. As in Z1–Z3, Zeke could have decided otherwise; but unlike in Z1–Z3, Zeke could not have made the decision to shoot Chad straightaway at a different time. There was only that one moment at which that decision—the decision to shoot Chad straightaway—could have been made. The decision could have gone otherwise, but the time of the decision could not have been different, at least insofar as this was up to Zeke (Of course, Chad could have brought it about that Zeke made his decision at a different time, by arriving at the window sooner or later than he actually did; but what Chad could have done differently won't account for Zeke's moral responsibility under PAP.). If it's dubious to make essential to moral responsibility a feature of an action that could have been otherwise but is arguably morally trivial, it's all the more dubious to do so when the feature is one that couldn't have been otherwise (as a result of the agent's doing) at all. At the very least, this is not a move one would expect in a defense of PAP!

We do not deny that it's possible to contrive cases in which a person may be morally responsible for A-ing at  $t$  while not being morally responsible for A-ing *simpliciter*, but such cases are unusual (hence the qualification above that it's only *typically* true that someone who is responsible for A-ing at  $t$  is also responsible for A-ing *simpliciter*). Considering such cases shows why Palmer cannot look to them for aid and comfort. Here's a slight variation on a well-known example from van Inwagen (1983, pp. 176–177). Ryder is astride a runaway steed, Dobbin, who can be turned but not stopped. There is one fork in the road, but both branches lead to Rome. Ryder may be responsible for taking the left fork or the right fork, but he isn't responsible for arriving in Rome—that isn't something that he in any sense *does*. Now add that one of the routes leading away from the fork is longer than the other one, and that Ryder knows this. Then Ryder isn't responsible for arriving in Rome, but (assuming he knows how long it will take to reach Rome via each road), he may well be responsible for arriving in Rome at  $t$  or at  $t + n$ . For an example in which a person is morally responsible for *deciding* to A at  $t$  but not for deciding to A, consider poor Patsy, who has a Martian chip implanted in her brain that the Martians are going to activate at  $t$ , causing her to decide to set off a bomb at the climax of a large public event. Just in time, a member of the anti-Martian resistance alerts her to the danger and gives her a lead-lined helmet that will slow down (by several minutes!) the signals sent to the chip in her brain. Then everyone (supporters and opponents of PAP alike) would agree that Patsy isn't morally responsible for deciding to set off the bomb *simpliciter*, because that's something the Martians did

to her; but she could be morally responsible for deciding to set it off at  $t$  instead of several minutes after  $t$ , when the crowd at the event will have dispersed, since she could have delayed the decision by donning the helmet.

What is striking is the *contrast* between such cases and typical cases of avoidable action, including those typical cases that are rendered atypical (and their actions unavoidable) by the introduction of a Frankfurt intervener. Compare the Martian case with *Tax Evasion 2*. Palmer does nothing to challenge Pereboom's claim that Joe is unable to decide otherwise; indeed, Palmer concedes this claim. Instead, he draws attention to the fact that Joe could have refrained from deciding *then*, because he could have decided (somewhat) *later*. But that's exactly how matters stand in the Martian case. Patsy cannot decide otherwise, but she *can* decide later rather than sooner. But surely there is a huge difference between these cases. One wants to say that Patsy is *not* morally responsible for deciding to set off the bomb, whereas Joe *is* morally responsible for deciding to evade taxes. But Palmer cannot say this. He has to treat the two cases as on a par, since Joe and Patsy both exhibit what is for Palmer the primary form of moral responsibility: responsibility for deciding *at t*. The critic of PAP, who looks to features of the actual sequence rather than the availability of alternative possibilities to ground moral responsibility, can readily account for the evident difference between the two cases; Palmer cannot.

Despite the fact that Palmer's thesis about the timing of decisions seems forced when applied to all but a handful of unrepresentative cases, he claims that it's Pereboom who bears the burden of showing why the agent in buffer cases is morally responsible for deciding *simpliciter*. He speculates that a defender of Pereboom's position might rely on something like the following inference (p. 268):

If a person is morally responsible for  $x$ , and  $x$  entails  $y$ , then he is morally responsible for  $y$ .

This inference is of course invalid, as Palmer points out (Example: Zeke's responsibility for killing Chad doesn't make him responsible for Chad's being mortal.). Nor, we might add, is the following more specific instance of this inference valid:

If a person is morally responsible for A-ing at  $t$ , and A-ing at  $t$  entails A-ing, then he is morally responsible for A-ing.

Ryder and Patsy provide counterexamples to this inference. But Pereboom needn't rely on any such inference, for his position is supported by the many cases in which the specific-instant timing of a decision is *prima facie* morally trivial or even outside the agent's control, as we showed above. In *typical* cases of moral responsibility, we suggest,

A person's moral responsibility for A-ing at  $t$  consists in his A-ing and in his being morally responsible for A-ing, where A occurs at  $t$ .

So Zeke's moral responsibility for deciding to shoot Chad at  $t$  consists in his deciding to shoot Chad and being morally responsible for deciding to shoot Chad, where this decision takes place at  $t$ . (Likewise, his responsibility for deciding to shoot Chad *from exactly where he stands behind the desk* consists in his deciding to

shoot Chad and being responsible for so deciding, where the decision occurs at a particular location behind the desk.) While it's also possible to come up with exceptional cases of moral responsibility in which the exact timing of the decision is salient and the agent is not morally responsible for deciding *simpliciter*, the buffer strategy does not rest on cases like these: neither *Tax Evasion 2* nor Hunt's *Revenge* contains anything like the special features that render Ryder and Patsy responsible only for the timing of their actions.

Summing up, we have argued that Palmer's response to the buffering strategy is unsuccessful, failing as it does in one of its two essential tasks. To recall, Palmer must show both that Joe the tax evader isn't morally responsible for what he cannot avoid (or, alternatively, for what he can't avoid being morally responsible for), and that what Joe *is* morally responsible for is relevantly avoidable—that Joe has a “robust” alternative to it. Even granting for argument's sake that paying more attention to his moral reasons at *t1* constitutes a robust alternative (a matter we take up in the next section), Palmer hasn't motivated his denial that Joe is morally responsible for deciding to evade taxes. In short, denying Joe's responsibility for deciding to evade taxes *simpliciter* seems no more plausible than denying Zeke's responsibility for deciding to shoot Chad *simpliciter* in Z1–Z3. Finally, since Palmer accepts that deciding to evade taxes *simpliciter* is unavoidable in *Tax Evasion 2*, this case remains a viable counterexample to PAP. We now turn to Franklin.

#### 4 Franklin's objection to Hunt

Like Palmer, Franklin is committed to restricting the content of agents' moral responsibility in buffer cases to decisions at particular times; as will be seen, however, Franklin is less explicit about the restriction. Our focus will be on Franklin's argument that the agent retains a robust alternative possibility in cases like *Revenge* and *Tax Evasion 2*. While this argument appears to be independent of the temporal restriction, we will show that this is not the case; without this restriction, the buffering strategist can reasonably maintain that the agents lack robust alternatives on a plausible construal of Pereboom's criterion for robustness.

To begin, Franklin sets out to defend a more precise variant of PAP, one that concerns responsibility for action at a specific time:

PAP<sub>1</sub>: Necessarily, an agent S is morally responsible for  $\varphi$ -ing at  $t_2$  only if S (i) was able at  $t_2$  to avoid responsibility for  $\varphi$ -ing at  $t_2$  or (ii) if S is unable at  $t_2$  to avoid responsibility for  $\varphi$ -ing at  $t_2$ , then there is some earlier time  $t_1$  at which S  $\psi$ -ed (where  $\psi$  might be an action or omission) that explains why S is unable at  $t_2$  to avoid responsibility for  $\varphi$ -ing at  $t_2$  and S was able at  $t_1$  to avoid responsibility for  $\psi$ -ing at  $t_1$  (2011, p. 192).

Insofar as PAP<sub>1</sub> pertains to responsibility for a temporally specific occurrence, it may be natural to suppose that Franklin believes that agents in buffer cases are morally responsible only for such temporally specific occurrences. However, Franklin doesn't actually say this; nor does he give any indication that he sees his



argument for the robustness of these agents' alternative possibilities as depending on this temporal restriction. This will prove to be an important issue.

We shall focus here on (i), leaving (ii), the "tracing" condition, to one side. The first thing to notice is that, even if PAP<sub>1</sub> can be defended, this won't by itself help Frankfurt's opponents. After all, suppose that Jones in *Revenge* is morally responsible for deciding to kill Smith, but *not* for so deciding at the precise moment he does. And suppose that Jones is unable to avoid deciding to kill Smith *simpliciter*, something both Franklin and Palmer allow. In that case, there will be a single state of affairs—that Jones has decided to kill Smith—that is both unavoidable and such that Jones is morally responsible for it.<sup>11</sup> That another state of affairs *is* avoidable—that Jones decides at  $t_1$  to kill Smith—will be cold comfort to defenders of the traditional association of responsibility with avoidability. Since deciding to kill Smith is *ex hypothesi* unavoidable in Jones's circumstances, preserving the traditional association will require Franklin not only to defend PAP<sub>1</sub>, but also to suppose that *all* Jones is morally responsible for is deciding at a precise time to kill Smith (and not for so deciding *simpliciter*). If this supposition is untenable, as was suggested in the previous section, Franklin's defense of PAP<sub>1</sub> is at best a Pyrrhic victory.<sup>12</sup>

As it turns out, however, Franklin needs the temporal restriction for a second reason as well; for without it, his defense of PAP<sub>1</sub> itself is implausible. Or so we shall now argue.

Franklin aims to show that both the truth and the *relevance* of PAP<sub>1</sub> withstand the challenge from buffer cases; such a principle is relevant, the idea is, just in case the alternative possibility can plausibly be thought to ground or explain the agent's responsibility—just in case it is robust. Like Palmer, Franklin (p. 196, note 16) adopts Pereboom's condition for robustness (at least for discussion's sake), an "epistemic condition for relevance" that he says is also "suggested" by Hunt.<sup>13</sup> Franklin formulates the condition as follows:

<sup>11</sup> More carefully, Jones will be both morally responsible for deciding to kill Smith, and *unable to avoid moral responsibility* for so deciding. The crucial point here is that the *mere possibility* of avoiding responsibility for this decision won't satisfy Pereboom's (or any plausible) condition of robustness. So, while it is indeed *possible* that Jones avoids responsibility for this decision by triggering Black's intervention, this possibility is not, as Franklin puts it, sufficient to preserve the *relevance* of PAP<sub>1</sub>. More about this shortly.

<sup>12</sup> Notice that we are *not* arguing that PAP<sub>1</sub> fails simply because PAP does, or more generally that if some version of PAP fails, none is tenable. Rather, we are pointing out that promising variants of PAP must meet certain conditions, namely, they must show either that no item for which the agent is morally responsible (or blameworthy) is unavoidable, *or* that no item for which the agent is morally responsible (blameworthy) is such that the agent's *responsibility* (blameworthiness) for it is unavoidable. Our objection to PAP<sub>1</sub> is that it doesn't meet this disjunctive condition. Thanks to anonymous referee for prompting this clarification.

<sup>13</sup> Franklin (p. 195). Franklin does not say where this condition is to be found in Hunt's paper. The condition Hunt actually invokes, as we shall explain in Sect. 5, is stronger than Pereboom's epistemic condition, though it involves an epistemic component. On the following page (note 17), Franklin cites a Q-and-A session at which Hunt appeared to endorse the epistemic condition as a minimal necessary condition for the robustness of an alternative, but this is perfectly consistent with Hunt's advocating stronger conditions as well, which he does.



- (EC) An ability to do otherwise is relevant to explaining why an agent is morally responsible for  $\varphi$ -ing only if the agent understands that in exercising this ability he would thereby avoid moral responsibility for  $\varphi$ -ing<sup>14</sup> (p. 195).

As we have seen, Pereboom maintains that Joe's decision in *Tax Evasion 2* doesn't meet this condition because Joe doesn't understand that he would avoid moral responsibility for deciding to evade taxes if he were to pay more attention to his moral reasons. And the same thing goes for Jones in *Revenge*, the case on which Franklin focuses: Jones doesn't understand that he would avoid moral responsibility for deciding to kill Smith if he were to consider not killing Smith.

In response, Franklin contends that EC admits of two possible readings, and that Jones satisfies the condition on the reading that should be accepted. (If he is right, the same will hold for *Tax Evasion 2*.)

On one reading (p. 198), an alternative possibility counts as robust provided that the agent understands that, by realizing that alternative possibility, she will avoid being morally responsible for her action *at the time at which she is actually responsible for it*. Jones clearly meets this condition, Franklin observes; for Jones cannot fail to understand that, by considering not killing Smith at  $t$ , he will be postponing his decision, and thereby avoiding responsibility for deciding at  $t$  to kill Smith. On the second, stronger reading (*ibid.*), an agent must understand that, by realizing the alternative possibility, she will avoid being responsible for her action *then and at all later times*. It seems clear that Jones doesn't meet this requirement; for all he knows, he might well decide to kill Smith immediately after considering (at  $t$ ) not killing Smith.

We agree with Franklin that EC seems implausibly strong on the second reading. Whatever else, having a robust alternative possibility doesn't require an agent to understand that, if she realizes that possibility, she will *never* be morally responsible for a decision like the one she actually makes. If these were the only possible interpretations of EC, we would either have to accept the first, weaker reading or else reject EC altogether. However, these are not the only interpretations; many possibilities fall between these extremes, and hence (*pace* Franklin) the buffer strategist can pass between the horns of Franklin's "dilemma" (p. 199).

Consider this intermediate reading:

- (IR) An alternative possibility is robust only if the agent understands that, by realizing that possibility, she will be precluded from responsibility for taking the opportunity at hand to act as she does.

IR is weaker than Franklin's stronger interpretation, in that it doesn't require the agent to understand that she will *never* make such a decision. Thus IR doesn't require Jones to understand that he won't decide to kill Smith *the next time* their

<sup>14</sup> Here is Pereboom's recent statement of the condition (2005, p. 230):

*Robustness.* For an alternative possibility to be relevant per se to explaining an agent's moral responsibility for an action it must satisfy the following characterization: she could have willed something other than what she actually willed such that she understood that by willing it she would thereby have been precluded from the moral responsibility she actually has for the action.

Pereboom (forthcoming) further refines this condition; however, the details won't concern us here.

paths cross (or Joe to understand that he won't decide to cheat on his tax returns the next time he files them), only that he won't so decide this time around. Even so, Jones's decision to kill Smith doesn't meet this condition, for he doesn't understand that considering not killing Smith would preclude him from responsibility for taking his present opportunity to decide to kill Smith (for all he knows, he will exploit this opportunity in the next instant).

To be sure, IR's reference to "the possibility at hand" is vague. We believe that this vagueness matches our ordinary thinking about opportunities for action. For our purposes, the important point is that it's highly plausible to individuate present opportunities for action in such a way that these opportunities aren't usually instantaneous.<sup>15</sup> It's plausible to think, for example, that in a counterfactual scenario without Black, Jones's present opportunity to kill Smith would have ended when Smith left the party.

IR illustrates that Franklin's strong and weak readings of EC aren't the only options.<sup>16</sup> How does IR measure up against the weaker interpretation of EC, which Franklin endorses? We believe that this depends on how one understands what Jones is morally responsible for. If he is morally responsible only for deciding to kill Smith at *t*, and not for deciding to kill Smith *simpliciter*, then (assuming that some version of EC is correct) it is indeed plausible to think that Jones has a robust alternative possibility in virtue of understanding that considering (at *t*) not killing Smith would preclude him from responsibility for deciding at *t* to kill Smith. For then it would be up to him whether or not he is responsible for deciding at *t* to kill Smith.

On the other hand, if Jones is morally responsible for deciding to kill Smith, it's hard to see what turns on his understanding that he can avoid responsibility for so deciding at *t*. If he is indeed morally responsible for so deciding *simpliciter*, then considering not killing Smith *in Black's absence* before deciding on his own to kill Smith wouldn't absolve him of responsibility for so deciding. Why think, then, that what grounds or explains his responsibility for deciding to kill Smith in Black's presence is his understanding that he can avoid responsibility for so deciding *just then*? Understanding that he can avoid responsibility for so deciding *then* doesn't make his decision avoidable in any interesting sense.

By contrast, Jones's correctly understanding *in Black's absence* that he is able to consider not killing Smith, and then (potentially) decide *not* to kill Smith, *does* seem to satisfy an interesting avoidability requirement, one that might plausibly be thought to ground his responsibility for deciding to kill Smith. As *Revenge* seems to show, however, Jones can be morally responsible for deciding to kill Smith even though, given Black's presence, he doesn't satisfy this requirement.

What emerges, then, is that EC is promising on the weak reading only if we suppose that Jones's moral responsibility is temporally restricted. If Jones is morally

<sup>15</sup> One reason for denying that such opportunities are instantaneous is precisely that we have limited control over the exact timing of our decisions (cf. Pereboom 2001, p. 32).

<sup>16</sup> Another plausible, intermediate option for the buffering strategist is this: "An alternative possibility is robust only if the agent understands that, by realizing that possibility, she will be precluded from responsibility for her action *barring an unforeseen change in her reasons for action*." And of course, combining this condition with IR would yield yet another intermediate reading.

responsible for deciding as he does *simpliciter*, IR seems clearly preferable to the weak reading. We argued in the previous section that the temporal restriction is implausible. Since Franklin's argument that buffer cases preserve robust alternative possibilities is unconvincing without this restriction, his defense of PAP<sub>1</sub> fails; he has not shown that the truth and relevance of this principle withstand buffer cases. But even if the robustness of Jones's alternative possibility could be defended without this restriction (so that, for all the buffering strategist has shown, PAP<sub>1</sub> remains both true and relevant), the problem noted at the beginning of this section would remain. If Jones is morally responsible for deciding as he does *simpliciter* despite being unable to avoid responsibility for this, the truth of PAP<sub>1</sub> will be cold comfort to defenders of the traditional association of moral responsibility with avoidability.<sup>17</sup>

Summing up, then, if the temporal restriction is granted, there is no reason to prefer IR to the weak reading (or vice versa). But the temporal restriction should not be granted, and Franklin's defense of PAP<sub>1</sub> fails without it. Before concluding our discussion of Franklin, we shall consider a move that might be suggested on Franklin's behalf.

To repeat, Franklin, unlike Palmer, is not explicit about whether he takes the responsibility of agents in buffered Frankfurt cases to be temporally restricted, much less whether he takes it to be restricted to the utterly specific times at which their decisions occur. As noted, whether Franklin takes responsibility attributions to be temporally restricted is a separate question from whether PAP<sub>1</sub> pertains to responsibility for actions performed at specific times (as it explicitly does). Now in maintaining as we have that Franklin's defense of PAP<sub>1</sub> requires him to restrict responsibility to decisions-at-specific-times, we have been supposing that the times involved will be *utterly* specific, as Palmer takes them to be. This supposition is natural in light of the fact that PAP<sub>1</sub> seems to use '*t*<sub>2</sub>' to denote the precise time when the action occurs: "...only if [the agent]...was able at *t*<sub>2</sub> to avoid responsibility for  $\varphi$ -ing at *t*<sub>2</sub>." Where the value of ' $\varphi$ ' is a *simple* mental act, '*t*<sub>2</sub>' will denote the utterly specific instant at which  $\varphi$  occurs. In light of Z1–Z3, we have argued, it's implausible to suppose that we are morally responsible *only* for making the decisions we do at the utterly specific times at which we make them. If Franklin's defense of PAP<sub>1</sub> requires this restriction, it is unpersuasive.

It might be asked, however, whether Franklin can construe the timing restriction more liberally than Palmer does, and, if so, whether this will help his defense of PAP<sub>1</sub>. Suppose that Franklin were to say that '*t*<sub>2</sub>' stands not for the utterly specific instant at which Jones's decision to kill Smith occurs in the actual sequence, but for some very short interval (a half second, say) that *includes* that utterly specific

<sup>17</sup> Another concern about Franklin's argument is this. Recall Franklin's reason for believing that the strong reading of EC is too strong: it's implausible to require that someone understand that she will *never* be morally responsible for performing such an action. Notice, however, that this requirement no longer seems implausibly strong if we adopt the temporal restriction. For in that case, Jones need only understand that, if at *t* he were to consider not killing Smith, he would *never* be morally responsible for deciding *at t* to kill Smith! At the least, addressing this concern would require Franklin to re-characterize his opponent's situation if Franklin insisted upon the temporal restriction.

instant. Then Franklin's claim would be that Jones understands that by considering (during this interval) not killing Smith, he will avoid being morally responsible for deciding at any time during this interval to kill Smith; and so PAP<sub>1</sub> survives.

We do not find this move promising. Just as it's implausible (as Z1–Z3 illustrate) to think that we are typically morally responsible only for deciding as we do at some utterly specific instant, so it's implausible to think that we are typically morally responsible only for deciding as we do within some very brief interval. It's far more plausible to think that Jones, like Zeke, is morally responsible for deciding as he does *simpliciter*. Nor is anything gained by lengthening the interval. To see this, let  $t_2$  span the length of time (however long it turns out to be) that Smith remains in Jones's line of sight at the party. Jones may well understand that if he were (*per impossibile*, given Black's presence) to consider not killing Smith throughout that interval (i.e. until Smith exits his line of sight), he would not decide to kill Smith during  $t_2$ . But he does not understand that, by considering not killing Smith at *some* point while Smith remains in his line of sight, he will avoid responsibility for deciding upon the killing at *any* time when Smith remains in his light of sight. For he does not know that considering not killing Smith during this interval will trigger Black's intervention; and for all he knows, he remains free to decide to kill Smith (at some later point) during this interval, after he has considered not killing him. Thus Pereboom's requirement for robustness isn't met.<sup>18</sup>

## 5 Robustness revisited

We have adopted EC, Pereboom's requirement for robustness, because it figures in both Palmer's and Franklin's responses to the buffering strategy. We believe that this requirement helpfully illuminates the important epistemic dimension of robustness; however, a closer examination of EC is beyond the scope of this article. In this section, we call attention to another plausible test for robustness, one that has been suggested independently by Hunt (2005) and Shabo (2007). Unlike Pereboom, our aim in presenting this test is not to identify a necessary condition for robustness, but rather to underscore the burden of proof facing those who claim that a particular alternative possibility plausibly grounds an agent's responsibility for her action.

Anticipating the objection that Jones in *Revenge* has a robust alternative—namely, he is able to *consider* the possibility of not killing Smith, thereby triggering Black's intervention—Hunt writes:

If, as in *Revenge*, an agent is morally responsible in virtue of being *blameworthy*, a PAP-relevant alternative should be such that, by accessing it,

<sup>18</sup> Suppose that Black somehow knows that Smith will move at  $t_3$ ; and suppose that Black will—at the last possible instant before  $t_3$ —impel Jones to decide to kill Smith if Jones hasn't already decided by then to kill him. We believe that Jones is morally responsible for deciding to kill Smith *simpliciter*, as well as for deciding to kill Smith by  $t_3$ , notwithstanding that both of these things are unavoidable (cf. Pereboom 2001, pp. 30–32, and forthcoming).

the agent would avoid (or mitigate) blame; but merely considering the possibility of not killing Smith would have no effect by itself on Jones's blameworthiness (2005, p. 138).

We believe that Hunt's reply to this objection, suitably elaborated, has broad application.

To begin, grant that any indeterministic Frankfurt case will involve some alternative possibility, some room for departure from the actual course of events. Now suppose that a PAP-defender claims that some such alternative possibility is robust; in claiming this, she is claiming that the agent is morally responsible *in virtue of* this alternative possibility. Put another way, she holds that the presence of this particular alternative possibility makes a crucial difference such that without it, the agent wouldn't have been morally responsible for the action, even if the action were causally undetermined (so that some alternative possibility remained).

It's easy to see that many alternative possibilities don't meet this condition. Cases in point include Jones's suddenly blacking out just before the time of his decision, and his forming an intention to abandon his revenge plan due to a random glitch in his central nervous system. Even if these outcomes are consistent with Jones's immediate past and the laws of nature, he surely isn't morally responsible in virtue of them. Or consider some of the cases that would be available on the temporal restriction favored by Palmer and Franklin, under which any action inconsistent with Jones's deciding *at that time* to kill Smith (and known by Jones to be inconsistent with his so deciding then) would count as a robust alternative possibility. Perhaps Jones can walk and chew gum at the same time, but he probably can't add  $2,492 + 759$  and decide to kill Smith at the same time, or stick a fork in his left eye and decide to kill Smith at the same time, and he probably also knows that he can't do these things simultaneously. But these surely aren't the sorts of alternatives that explain Jones's moral responsibility for deciding at *t* to kill Smith.

What about Jones's considering not killing Smith? Unlike blacking out and suffering a random neurological mishap, this occurrence *is* up to Jones, and unlike adding  $2,492 + 759$  and sticking a fork in his left eye, this action seems relevantly related to the action for which Jones is being judged morally responsible. Thus it's at least a candidate for being robust.

To see why it's an unsuccessful candidate, Hunt suggests, we should ask what effect realizing this alternative possibility would have on Jones's moral responsibility for deciding to kill Smith. More specifically, Hunt asks, insofar as Jones is morally blameworthy for so deciding in the actual sequence of events, what effect (if any) would his considering not killing Smith *by itself* have on Jones's blameworthiness? Would Jones thereby avoid (or mitigate) his blame? Hunt believes not, and that this creates a presumption against the view that Jones is morally responsible in virtue of this possibility.

We shall look first at how this thought is fleshed out in Hunt (2005), and then propose a further way of developing it. As *Revenge* is set up, considering not killing Smith is the (necessary) first step Jones can take toward deciding otherwise. He is certainly blameworthy for not taking that step (just as Joe is blameworthy for not attending to his moral reasons against tax evasion). But Hunt doesn't claim that

Jones's blameworthiness for *not even considering not killing Smith* violates PAP; rather, his claim is that Jones's blameworthiness for *deciding to kill Smith* violates PAP. The alternative possibility of considering not killing Smith could plausibly be thought to explain Jones's moral responsibility for the former; the question is whether it can plausibly explain his moral responsibility for the latter. To shed light on this question, Hunt looks at a case in which an agent's refraining from taking a first necessary step toward not A-ing might plausibly explain his moral responsibility for A-ing, and might explain it even if a counterfactual intervener were to ensure that this first step would be the last. In *Smoker*,

Jack realizes that his first-order desire for cigarettes stands a good chance of yielding to his second-order desire to quit smoking if (but only if) he joins a 12-step program, while Mack, a smoking buddy who fears he will be left to smoke alone if Jack quits, stands ready, should Jack join the program, to slip him a pill which would make his craving for nicotine irresistible. Then Jack cannot access a decision to quit smoking, but he can (moved by the better angels of his second-order desire) decide to join the 12-step program. A defender of PAP might plausibly appeal to this alternative's availability to explain Jack's moral responsibility in the case where he doesn't join the program and continues to smoke "on his own," without the ministrations of Mack's pill (*ibid.*).

Hunt suggests that what makes this alternative robust with respect to Jack's responsibility for continuing to smoke is that Jack would choose it as a means toward the end of quitting; in joining the 12-step program, Jack would be doing what he can to quit; he would be *trying*. This condition on robustness is partly epistemic: if Jack thinks that 12-step programs are dance schools, his joining one wouldn't constitute a PAP-relevant alternative. But it's more importantly a reasons condition: if he joins the program in order to meet women, with no thought of quitting, he has not thereby accessed an alternative that would exempt him from or mitigate his blame under PAP.

Return now to the claim that Jones's considering not killing Smith is a PAP-relevant alternative to Jones's deciding to kill Smith. Does it satisfy the condition for robustness that Jack's joining the 12-step program does in *Smoker*? Hunt thinks not, and for a very simple reason. Jones's considering not killing Smith is robust, on the reasons condition suggested by *Smoker*, only if Jones considers not killing Smith *with the end of not killing Smith in mind*, that is, with the *understanding* that this is a step in the direction of not killing Smith, and with the *intention* of taking that step in order to reach that end. But then Jones's consideration of not killing Smith (call this Consideration 1), if robust, must be preceded by the thought of not killing Smith; because it is only insofar as it is directed by this prior thought that Consideration 1 satisfies Hunt's reasons condition for robustness. But then this prior thought of not killing Smith (call it Consideration 2) will trigger the intervener.<sup>19</sup> If Hunt's reasons

<sup>19</sup> Note that "considering," as Hunt uses the term, is a simple mental act; in the case of Jones's considering not killing Smith, it is, "minimally, the conscious awareness, however brief, of *not killing Smith* as a possibility for him" (p. 133). Jones cannot take not killing Smith as the goal of an activity without being consciously aware of not killing Smith as a possibility for him. It's in this sense that considering not killing Smith, as a means toward the goal of not killing Smith, presupposes a prior consideration of not killing Smith.



condition is correct, then, the only way to make Jones's considering not killing Smith a robust alternative is to put it out of Jones's reach, buffered by a prior consideration of not killing Smith.

This may be cutting things pretty thin, psychologically speaking; it also assumes that Hunt's reasons condition is acceptable, an assumption we find plausible but in need of further investigation. We therefore propose a second way of fleshing out the essential idea that Jones's moral responsibility for deciding to kill Smith cannot rest on the possibility of his considering not killing Smith, because his considering this would not *by itself* result in the avoidance or mitigation of his blame.

Suppose that Black is absent from the scenario, and that Jones ends up deciding on his own to kill Smith, but that before so deciding, he pauses to consider not killing Smith. If considering not killing Smith *by itself* eliminated or mitigated Jones's blameworthiness for deciding to kill Smith, Jones would at least be less blameworthy than he is when he decides to kill Smith straightaway. Clearly, however, pausing to consider not killing Smith doesn't make this difference; *this* departure from the actual course of events wouldn't make Jones any less blameworthy for deciding to kill Smith.

Now bring back Black. In the actual sequence, where Jones decides on his own at  $t$  to kill Smith, Jones has this bit of unexploited "leeway" available: he is able to consider not killing Smith instead of deciding at that moment (as he actually does) to kill Smith. Our PAP-defender maintains that this unexploited leeway is the key to Jones's moral responsibility for his decision to kill Smith in the actual sequence. If this is right, however, it seems reasonable to ask what would happen if Jones exploited this leeway in Black's absence. If considering not killing Smith and then deciding to kill him anyway would not make Jones less blameworthy than when he kills Smith straightaway, it's hard to see why so much should hinge on this unexploited leeway. Since exploiting this alternative possibility in Black's absence would not eliminate (or mitigate) Jones's blameworthiness for deciding as he does, this unexploited alternative doesn't plausibly ground Jones's blameworthiness in the actual sequence in *Revenge*.

More generally, if a PAP-defender claims to identify a robust alternative possibility in a Frankfurt case, we believe that it's fair to ask her whether, in Black's absence, exploiting that alternative possibility before adopting the course she actually does adopt would eliminate (or mitigate) her blameworthiness. If the answer is "no," it is especially incumbent on her to explain why that unutilized "possibility space" should make the crucial difference to the agent's responsibility (Shabo 2007, p. 71).

Perhaps an adequate explanation is available in some cases. For instance, if we remove the psychological buffer from *Revenge*, so that Jones was able to decide not to kill Smith exactly when he *actually* decided to kill him, it's plausible to think that Jones had a robust alternative to deciding to kill Smith. Moreover, it's plausible to think this even supposing that Black would have impelled Jones to decide to kill Smith a moment later. Yet in Black's absence, if Jones had decided against killing Smith, only to change his mind very shortly thereafter and decide to kill Smith after all (say, because he was suddenly struck by a powerful reason for killing Smith that he had previously overlooked), Jones would arguably be just as blameworthy as he



would be for deciding straightaway to kill Smith. So, Jones's unexploited leeway in Black's presence arguably constitutes a robust alternative possibility, notwithstanding that exploiting this leeway in Black's absence would not by itself mitigate or eliminate his blameworthiness for his decision.<sup>20</sup>

Now if we had maintained that an affirmative answer to our question—"Would exploiting this leeway in Black's absence have mitigated or eliminated Jones's actual blameworthiness for his decision?"—is a necessary condition for robustness, there might well be an objection here. However, we have not maintained this. Instead, we have maintained that *it's incumbent on our opponent to explain why* an alternative possibility should be considered robust when the answer to this question is negative. Perhaps there is a good answer to this question in the case of decisions. Perhaps, that is, there is room to explain why being able to decide against killing Smith should be considered a robust alternative, even if exploiting this possibility in Black's absence before deciding to kill Smith after all would not lessen Jones's blameworthiness. Be this as it may, we believe that the question is a good one, and that there is a presumption against any view that cannot give it an affirmative answer. If exploiting the residual leeway in Black's absence would not lessen Jones's blameworthiness, why think that his actual responsibility in Black's presence hinges on this unexploited leeway?

According to both Palmer and Franklin, the allegedly robust alternative possibility in buffered Frankfurt cases consists not in the agent's avoiding *deciding* at a particular time to do something, but rather in avoiding *responsibility for* so deciding.<sup>21</sup> Yet the above question (or a close variant of it) remains applicable; after all, not all ways of avoiding responsibility by acting under Black's impetus constitute robust alternative possibilities. If a random neurological malfunction (or some other involuntary occurrence) is what triggers Black's intervention and thereby absolves Jones of responsibility, we wouldn't count this alternative possibility as robust. We can reasonably ask, then, whether avoiding responsibility in the relevant way in Black's absence would *by itself* eliminate (or mitigate) the agent's blameworthiness for deciding as he does. Intuitively, the answer still seems to be "no": the agent would be no less blameworthy for deciding as he does after considering not killing Smith (*Revenge*) or attending to his moral reasons (*Tax Evasion 2*) than if he so decided straightaway. For this reason, it remains hard to see why Jones's responsibility in Black's presence should hinge on this unexploited leeway.

Of course, things look different if we suppose that the agent is morally responsible *only* for deciding as he does at a precise time. In that case, in Black's

<sup>20</sup> Thanks to an anonymous referee for prompting us to address this point.

<sup>21</sup> In correspondence, Franklin has denied that he takes the robust alternative possibility in buffered Frankfurt cases to consist in being able to *avoid responsibility for* the action, as opposed to avoiding the action. Since, however, his PAP<sub>1</sub> is stated in terms of being able to avoid responsibility for an action, and since, throughout his article, he supports his claims about agents having robust alternatives by appeal to what they can avoid responsibility for, we believe that this is a reasonable reading. He writes: "PAP<sub>1</sub> makes it clear that what is at stake in principles like PAP is the ability to avoid responsibility for some event rather than being able to bring about some alternative event. What is crucial for moral responsibility is not that I am able to prevent the action from occurring, but that I am able to avoid responsibility for its occurring" (p. 192).

absence, the agent would indeed avoid blameworthiness for that particular decision by simply postponing the decision, even if he would be no less blameworthy for the later decision. Without this temporal restriction, we believe the onus is clearly on the PAP-defender to explain why the unexploited leeway in *Tax Evasion 2* and *Revenge* makes the crucial difference when exploiting it wouldn't by itself preclude (or mitigate) blameworthiness for the decision. We see no way to discharge this onus, and we have argued that the temporal restriction should be rejected. Quite apart from Pereboom's epistemic condition, then, there is good reason to believe that the agents in *Tax Evasion 2* and *Revenge* are morally responsible for deciding as they do, notwithstanding that they lack robust alternatives.

**Acknowledgments** We would like to thank Derk Pereboom, Chris Franklin, and an anonymous referee for *Philosophical Studies* for helpful comments that led to several improvements in and additions to the manuscript. David Hunt's work on this paper was made possible through the support of a grant from the John Templeton Foundation.

## References

- Blumenfeld, D. (1971). The principle of alternative possibilities. *Journal of Philosophy*, 68, 339–345.
- Fischer, J. M. (1994). *The metaphysics of free will: An essay on control*. Cambridge, MA: Blackwell Publishers.
- Frankfurt, H. (1969). Alternate possibilities and moral responsibility. *Journal of Philosophy*, 66, 829–839.
- Franklin, C. E. (2011). Neo-Frankfurtians and buffer cases: The new challenge to the principle of alternative possibilities. *Philosophical Studies*, 152, 189–207.
- Ginet, C. (1996). In defense of the principle of alternative possibilities: Why I don't find Frankfurt's arguments convincing. *Philosophical Perspectives*, 10, 403–417.
- Ginet, C. (2002). Review of Pereboom's *Living without free will*. *Journal of Ethics*, 6, 305–309.
- Hunt, D. (2000). Moral responsibility and unavoidable action. *Philosophical Studies*, 84, 195–227. Reprinted in (2005) J. M. Fischer & G. Watson (Eds.), *Free will: Critical concepts in philosophy*. London: Routledge.
- Hunt, D. (2005). Moral responsibility and buffered alternatives. *Midwest Studies in Philosophy*, 29, 126–145. Reprinted in (2005) J. M. Fischer & G. Watson (Eds.), *Free will: Critical concepts in philosophy* (London: Routledge).
- Kane, R. (1985). *Free will and values*. Albany: State University of New York.
- Kane, R. (1996). *The significance of free will*. New York: Oxford University Press.
- Kane, R. (2000). The dual regress of free will and the role of alternate possibilities. *Philosophical Perspectives*, 14, 57–79.
- Palmer, D. (2011). Pereboom on the Frankfurt cases. *Philosophical Studies*, 153, 261–272.
- Pereboom, D. (2000). Alternate possibilities and causal histories. *Philosophical Perspectives*, 14, 119–138.
- Pereboom, D. (2001). *Living without free will*. Cambridge: Cambridge University Press.
- Pereboom, D. (2005). Defending hard incompatibilism. *Midwest Studies in Philosophy*, 29, 228–247.
- Pereboom, D. (forthcoming, 2012). Frankfurt examples, derivative responsibility, and the timing objection. *Philosophical Issues*.
- Shabo, S. (2007). Flickers of freedom and modes of action: A reply to Timpe. *Philosophia*, 35, 63–74.
- Shabo, S. (2010). Uncompromising source incompatibilism. *Philosophy and Phenomenological Research*, 80, 349–383.
- Stump, E. (1996). Libertarian freedom and the principle of alternate possibilities. In J. Jordan & D. Howard-Snyder (Eds.), *Faith, freedom, and rationality* (pp. 73–88). Lanham, MD: Rowman and Littlefield.
- Stump, E. (1999). Dust, determinism, and Frankfurt: A reply to Goetz. *Faith and Philosophy*, 16, 413–422.
- van Inwagen, P. (1983). *An essay on free will*. Oxford: Clarendon Press.

- 
- Widerker, D. (1995). Libertarianism and Frankfurt's attack on the principle of alternative possibilities. *Philosophical Review*, *104*, 247–261.
- Widerker, D., & McKenna, M. (2003). *Moral responsibility and alternative possibilities: Essays on the importance of alternative possibilities*. Burlington, Vt: Ashgate.