



2007

Black the libertarian

David P. Hunt

Follow this and additional works at: <https://poetcommons.whittier.edu/phil>



Part of the [Philosophy Commons](#)

Black the Libertarian

David P. Hunt

Whittier College

The most serious challenge to Frankfurt-type counterexamples to the Principle of Alternate Possibilities (PAP) comes in the form of a dilemma: either the counterexample presupposes determinism, in which case it begs the question; or it does not presuppose determinism, in which case it fails to deliver on its promise to eliminate all alternatives that might plausibly be thought to satisfy PAP. I respond to this challenge with a counterexample in which *considering* an alternative course of action is a *necessary condition* for *deciding* to act otherwise, and the agent does not in fact consider the alternative. I call this a “buffer case,” because the morally relevant alternative is “buffered” by the requirement that the agent first consider the alternative. Suppose further that the agent’s considering an alternative action—entering the buffer zone—is what would trigger the counterfactual intervener. Then it would appear that PAP-relevant alternatives are out of reach.

I defend this counterexample to PAP against three objections: that considering an alternative is *itself* a morally relevant alternative; that buffer cases can be shown to contain *other* alternatives that arguably satisfy PAP; and that even if the agent’s *present* access to PAP-relevant alternatives were eliminated, PAP could still be satisfied in virtue of *earlier* alternatives. I conclude that alternative possibilities are a normal symptom, but not an essential constituent, of moral agency.

Keywords: alternatives, buffer (buffer case, buffer strategy), decision, free will, libertarian, moral responsibility, Principle of Alternate Possibilities.

In Harry Frankfurt’s famous counterexample to the Principle of Alternate Possibilities, or *PAP*—roughly, that one is morally responsible for an action only if one could have done otherwise—a “counterfactual intervener” named Black would have forced an agent named Jones to decide to murder a fellow named Smith if Jones hadn’t decided to murder Smith on his own, making Black’s intervention unnecessary. In such a situation, Frankfurt claimed, the agent seems morally responsible (Jones acted on his own without interference from Black), yet unable to do otherwise (Black was ready to nip any alternative in the bud). So PAP is false.¹

It is widely agreed that Frankfurt's original counterexample comes up short. Libertarian critics, in particular, have complained that Frankfurt's nefarious Black doesn't take the hapless Jones's agency seriously enough. Suppose Jones is indeed a libertarian agent. Now how exactly is Black to know when to intervene? "Black is an excellent judge of such things,"² Frankfurt blithely noted. But this won't do. Black's judgment, however excellent, must rest on some sign of how Jones will act. If the sign causally determines the indicated action, Frankfurt has begged the question against the libertarian defender of PAP. But if the sign does not causally determine the action, it's possible that the sign occur but Jones *not* act in the indicated way, in which case Black cannot reliably eliminate the possibility of Jones's acting otherwise. In either case, Frankfurt has failed to defeat PAP.

In response to this and other objections, a second generation of counterexamples has sprung up. There are three in particular that I have found sufficiently persuasive to defend in print. One of these replaces the counterfactual intervener in Frankfurt's thought-experiment with the theistic God. Arguably, Jones can't do otherwise if an infallible believer knew what he would do before he was even born. But we can also say of God's foreknowledge what Frankfurt said of Black's device: it "played no role at all in leading him to act as he did," "could have been subtracted from the situation without affecting what happened or why it happened in any way," is "irrelevant to the problem of accounting for [Jones's] action," and "does not help in any way to understand either what made him act as he did or what, in other circumstances, he might have done."³ So we should hold Jones morally responsible in both cases. For this counterexample to succeed, however, one must reject (as I do) the various subterfuges by which alternative possibilities have been defended against divine foreknowledge, *and* accept (as I also do) the conceptual coherence of an infallible foreknower. Since many philosophers would reject at least one of these, the divine foreknowledge scenario can't be expected to break the dialectical stalemate over PAP. Nor can this be expected from so-called blockage cases, which bring the walls of Locke's locked room right inside Jones's cranium.⁴ The idea here is to posit an actual sequence of mental states, which Smith traverses on his own, leading up to the decision to murder Smith; then to add the condition that, owing to a freakish coincidence, all alternative neural pathways happen to be non-functional. The worry about such cases is that a blockage that form-fits snugly around Jones's actual mental sequence may beg the question against indeterminism. I'm not sure that it does; but then I'm not sure that it doesn't.⁵

This brings me to the third case, on which I will focus in this essay. I will first present the case and point out the distinctive features that (I claim) equip it to evade the libertarian charge of begging the question. I will then identify and respond to what I take to be the three strongest grounds for holding that this kind of case does *not* successfully eliminate all PAP-relevant alternatives. Assuming that I deal with these three objections satisfactorily, this third case—which carries considerably

less baggage than the first two cases—would appear to pose a serious difficulty for defenders of PAP.

Buffer Cases

Suppose that Black himself is a libertarian and that he approaches his “Jones problem” with the full realization that he is dealing with a libertarianly free agent. Confronted with Frankfurt’s original scenario—call it *Assassin*—Black knows that Jones could thwart him. So instead of looking for a sign that can operate as advertised only in a deterministic world, Black waits . . . and gets a lucky break. He manages to catch Jones in a set of circumstances, and at a point in Jones’s response to those circumstances, that allow Black to act, despite his libertarian convictions. Call this new scenario, which I will now describe, *Revenge*.

Smith makes Jones the butt of a demeaning practical joke. The first thought that occurs to Jones, after realizing what Smith has done to him, is that he would like to kill Smith. He goes for a walk, escaping the immediate pressures of the situation and giving himself ample opportunity to pull back from this line of thinking. Given the kind of person Jones is, and given the situation in which he finds himself, the alternative of *not* killing Smith is not unthinkable for him; moreover, should he entertain this alternative, nothing would prevent him from deciding and acting on it. But Jones could decide (and act) otherwise only if he first considered acting otherwise, and he never does this (though he could); instead, he nurses his grievance without respite, while the idea of killing Smith becomes more and more attractive to him. Having finally decided to do the deed, he gets a gun, tracks Smith down, and shoots him dead.

Before introducing the counterfactual intervener, let us note some of the salient features of *Revenge* as it has been described up to this point. In the first place, the idea of killing Smith comes to Jones unbidden (so the question whether he can do otherwise in *this* respect—i.e., *not* have the idea instead of having it—does not even arise).⁶ In the second place, once the idea has come before his mind, it is open to Jones to embrace the idea, thereby deciding to kill Smith, and it is open to him to dismiss the idea, thereby deciding not to kill Smith (so nothing constrains him one way or the other). Thirdly, given Jones’s psychology and given the circumstances in which he finds himself, including the fact that he is already thinking about killing Smith, he can avoid deciding to kill Smith only if he first considers the alternative of not killing Smith (so his considering not killing Smith—minimally, the conscious awareness, however brief, of *not killing Smith* as a possibility for him—is a *necessary condition* for his accessing an alternative to his actual decision). Finally, Jones decides to kill Smith without ever considering the alternative of not killing Smith (so he makes his decision without deliberating). There is, I would suggest, no *prima facie* reason for regarding these features of *Revenge*, either jointly or severally, as incompatible with Jones’s moral responsibility.

The final element to be added to *Revenge* is the counterfactual intervener, which differs from the device in *Assassin* inasmuch as it is programmed to hijack Jones's mental processes and force him to decide to kill Smith if he so much as *considers* not killing Smith. With this device in place, there is no alternative to Jones's deciding to kill Smith: Jones can decide otherwise only if he first considers doing so, but then the device will force him to decide *to* kill Smith. So an alternative decision is not available to Jones in *Revenge*. Yet Jones, who in fact proceeds to murder Smith on his own, leaving the device untriggered, seems morally responsible for killing Smith. PAP is therefore wrong when it makes moral responsibility rest on the availability of morally relevant alternatives.⁷

Before continuing, we need to replace our vague gesture toward PAP with something more definite. Here's a formulation that is reasonably perspicuous:

PAP An agent X is morally responsible for performing an action A only if X did something B (which might or might not be identical to A) such that (i) X could have refrained from doing B and (ii) it is at least in part in virtue of X's having been able to refrain from doing B that X is morally responsible for performing A.

With this formulation in hand, we can now take stock. *Revenge* differs from *Assassin* in that it contains a "buffer zone" between the agent's actual pathway and any alternative pathways that might plausibly satisfy PAP. Entering the buffer zone triggers the intervener, so PAP-relevant alternatives are all out of reach. "Buffer cases" like *Revenge*, where intervention is triggered by a *necessary* condition for accessing a morally relevant alternative, appear to escape the usual objections to Frankfurt counterexamples like *Assassin*, which rest on the supposition that the prior signs to which the mechanism is responsive constitute *sufficient* conditions for the agent's performing or refraining from B (absent intervention by the mechanism).

Two kinds of objections can be brought against this "buffer strategy." One kind grants that *Revenge* eliminates all of Jones's morally relevant alternatives, while denying that Jones is then morally responsible for deciding to kill Smith; the other accepts Jones's moral responsibility while insisting that non-buffered alternatives remain. Because the second kind of objection seems to me to be the more serious of the two, it is the one that I will pursue in the remainder of this paper. In particular, I will look at three objections to the effect that, insofar as *Revenge* is a case in which Jones is genuinely responsible for his actions, it is a case in which alternatives sufficient to satisfy PAP are in fact available.⁸

The First Objection

In *Revenge*, as in *Assassin*, the counterfactual intervener cannot act quickly enough to squelch the triggering event itself. So this event—in *Revenge*, Jones's considering not killing Smith—remains an open possibility. But then Jones, who in actual-

ity does *not* consider not killing Smith, has available to him an alternative course of action that the pro-PAP party can exploit.

Reply

Whatever might be said of the residual alternatives in traditional Frankfurt counterexamples like *Assassin*, the alternative left open in *Revenge* does not satisfy clause (ii) of PAP. 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, 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.⁹

Compare this case with *Smoker*, in which 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 decide to join the 12-step program. A defender of PAP might with some plausibility 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 help of Mack's pill. But the appeal is plausible only on the assumption that Jack, in the alternative sequence of events, joins the program *as a means toward quitting*. (Contrast this with a case in which he joins the program in order to make the acquaintance of an attractive woman who attends the meetings: *this* alternative possibility, accessed for this reason, would not satisfy PAP.) In *Revenge*, however, undertaking to consider not killing Smith is not something that Jones does *as a necessary means* toward the goal of not killing Smith. Indeed, it would appear that undertaking to consider not killing Smith *could not* play this role for him, inasmuch as it would impose on him the incoherent task of considering not killing Smith (in order to take not killing Smith as his intended end) *before* undertaking to consider not killing Smith. (Alternatively and more coherently, Jones might engage in two, possibly continuous, considerings, the second of which would be morally significant because done as a means toward deciding not to kill Smith. In *Revenge*, however, this second considering won't be available to Jones, since the first PAP-irrelevant considering would trigger the intervener.)

We might give this objection more credit than it's due because we think Jones *should* have considered not killing Smith: we blame him, not only for deciding to kill Smith, but also for not even considering not killing him. Since we hold him responsible for this failure, and an alternative was available to him (nothing prevented Jones from considering not killing Smith), *Revenge* appears to contain at least one action that satisfies PAP.¹⁰ But this only satisfies the PAP-requirement for his *not considering not killing Smith*; it doesn't show PAP to be satisfied for his

deciding to kill Smith, nor does it appear that responsibility for the latter could simply piggyback on the former. The job of the Frankfurtian objector is to produce a single case in which PAP fails, not to show that it never succeeds.

This first objection is a generalized version of the typical first-blush negative reaction to Frankfurt-type counterexamples: that Jones could (and should) have *tried* to do/decide otherwise, and that if he does try and the device then intervenes, he has satisfied his moral obligations. *Revenge* is an improvement over *Assassin*, because Jones can't even try. *Revenge* rules out the possibility of trying (to decide) to refrain from murder, because one must consider doing X before trying to do X, allowing the device to intervene to prevent the trying before it can be initiated.¹¹

The Second Objection

It is simply not true that Jones's only immediate options, given that he is currently thinking about killing Smith, are either bringing this thought to fruition by deciding to kill Smith, or else replacing this thought with the alternative thought of not killing Smith. Jones might, for example, simply *fail* to decide to kill Smith, without ever considering not killing him; in this case the triggering event would not occur, and Jones would avoid killing Smith.

Reply

To count as an alternative satisfying PAP, Jones's refraining from deciding to kill Smith must be something that Jones *does*; it can't be something that is done to him, or that simply happens. It is the former kind of refraining for which Jones's considering not killing Smith serves in *Revenge* as a necessary condition. Now any case of Jones's not deciding to kill Smith will be a case of Jones's refraining (in this sense), or it won't be. If the latter—as when, e.g., Jones fails to decide to kill Smith because he is himself killed, or is overcome by sleep, or is distracted by a loud explosion (after which he forgets what he had been thinking about)—the alternative is of the wrong sort to satisfy PAP. If Jones is morally responsible for killing Smith, it isn't in virtue of *these* alternative possibilities, the actualization of which would simply make Jones the beneficiary of "moral luck." But if the alternative is a genuine case of refraining, it's not clear why we can't simply stipulate in setting up the example that Jones can move to this alternative from his present course of action only by first considering this alternative. Whether or not any actual agents are so constituted, it's hard to see why the supposition that Jones is so constituted would disqualify him as a moral agent. But then the mechanism can intervene to prevent Jones from accessing this alternative.

Perhaps the *does/done-to-him* distinction on which this initial reply rests is too artificial to cover all the alternatives the objector has in mind, making it possible to pass between the horns of the dilemma. Let us therefore consider one way things

might go that does not fit neatly into this rubric. It would seem that Jones could leave the decide-to-kill-Smith track, not just *via* a consideration of the possibility of not killing Smith, or under the prompting of an intervening cause, but in direct *reaction against* his contemplated course of action. Suppose, for example, he approaches Smith with gun drawn and suddenly realizes what it is he is about to do; he then flees in horror. Here he moves to the alternative pathway, not because he is attracted to it (he couldn't be attracted to it without first considering it, and this would trigger the intervener), but because he is repelled by what he is actually contemplating. There is some plausibility to the claim that he could *and should* have been affected in this way. The mechanism can certainly be reset to prevent Jones's reaching a contrary *decision* along this new route, but it can't intervene until Jones takes the first step along this route, and the fact that this first step is open to Jones seems morally relevant to his culpability for murder. At least it seems *more* relevant than the alternative of merely considering not killing Smith. For one thing, *feeling bad* about a prospective course of action represents a real if minimal departure from moral neutrality, whereas simply *entertaining an option* does not (by itself) commit the agent, even incipiently, to one side or the other. What's more, if Jones *were* to refrain from deciding to murder Smith, his discomfort at the thought of killing Smith might help explain why he didn't decide to kill Smith, whereas his having considered not killing Smith could not similarly explain why he opted for this outcome rather than the other.¹²

The problem with this reply is that we still need to understand how the availability of this overlooked alternative might satisfy PAP. It's hard to see how it could do so unless the alternative in question—feeling bad, entertaining a doubt, making a negative judgment—*ipso facto* put any decision to kill Smith in abeyance. But it obviously does no such thing: any negative reaction to Smith's murder on Jones's part might be outweighed in Jones's mind by other factors; or Jones's trajectory toward action might be akratically impervious to moral qualms, prudential calculations, and the like. PAP's defenders need an alternative that would allow the agent to avoid blame, and feeling bad or conflicted about what one is doing simply fails to do the job. So even assuming that Jones *could* react in the described way, acquiring a negative attitude toward his proposed action in one simple step, it's doubtful that this would rescue PAP. But one might also challenge the assumption. Let *Revenge* be fine-tuned so that Jones's responding in the suggested way would be completely out of character for him: someone who responded in this way just wouldn't be *Jones*. Perhaps Jones has killed before, on a number of occasions and in similar circumstances. He isn't such a hardened criminal that he *can't* view his actions from a moral point of view and feel an appropriate revulsion; but the idea that *this person* might be heading straight toward murder, looking neither to the right nor the left, and suddenly, in one simple step, undergo a PAP-relevant shift in attitude, seems like an appeal to *magic*. This is to expect too much of him, and the judgment that he is morally responsible for killing Smith surely does not rest on

this expectation being a reasonable one. This attempt to evade the dilemma posed in the first paragraph of this reply is a failure.

The Third Objection

If there are no PAP-relevant alternatives once Jones finds himself thinking about killing Smith, there may still be PAP-relevant alternatives available prior to the events described in *Revenge*. In particular, Jones's responsibility for killing Smith might derive from his responsibility for being a certain kind of person (e.g., the kind of person who could decide to kill someone, without even considering doing otherwise, just because he was humiliated by him), which in turn is grounded in earlier character-forming choices which themselves satisfy PAP.¹³

Reply

I believe that this objection fails on at least two counts. First, it simply challenges the critic of PAP to concoct settings for these earlier character-forming episodes that are themselves structurally similar to *Revenge*, in that a counterfactual intervener is present and the only PAP-relevant alternatives are buffered alternatives. Since the objector's appeal to ancient history is an apparent admission that *Revenge* itself contains no PAP-relevant alternatives accessible to Jones, it's not clear on what grounds the objector might insist that these earlier alternatives could not be rendered similarly inaccessible.

Call this extension of the buffer strategy into Jones's past *Global Revenge*. There is one significant difficulty with this response that needs to be addressed before moving on to the second reply to the objection.¹⁴ For Jones's earlier episodes of moral agency to conform to the pattern of *Revenge*, considering an alternative must always be a necessary condition for Jones to access that alternative, the device must be set to intervene if this necessary condition is satisfied, and the necessary condition must never be satisfied. This means that the Jones of *Global Revenge* apparently *never* considers an alternative to what he actually does, and such a person just seems too *weird* to count with any confidence as a moral agent: always acting on the first thought that pops into one's head is surely a sign of arrested development, neurological disorder, or other defect.¹⁵ But it's one thing to recognize a deficiency in this person's agency; it's another thing to say just what is missing.

In *Revenge*, Jones considers, wills, and then executes the tokening of a specific action-type. One way to understand the connection between *Revenge* and a putative character-forming episode is to suppose that there is some action-type A—one that the objector regards as morally perspicuous—which is tokened on both occasions. It's because of this similarity of type that Jones's responsibility on this earlier occasion can "transfer" to his decision in *Revenge*, despite his inability to access an alternative to this decision. Moreover (the objector might continue), this

earlier episode can be genuinely character-forming, entitling it to confer responsibility on later A-episodes, only if Jones, on that occasion, considered both A and not-A, and then decided and acted on his own, without the ministrations of the device. This suggests that the objector's real complaint against *Global Revenge* is something like this:

(G) There is no occasion on which Jones has considered doing A and considered not doing A and then reached his own decision about what to do, without interference from the device.

This is of course true in *Global Revenge*, since the elimination of PAP-relevant alternatives requires that either Jones's considering A or his considering not-A serve as a trigger for the device's intervention, and his considering both guarantees its intervention.

It may be possible to challenge (G) by distinguishing two senses in which one might consider a possible action prior to deciding to perform it. One sense is theoretical, which might be rendered this way:

(a) I might do A.

This is a simple recognition that A is among the things that it is possible for me to do. But a second sense is practical:

(b) Would that I might do A?

Here I consider A with an eye toward whether to do it. An affirmative answer to this question might be captured this way:

(c) Would that I might do A!

If (b) is the necessary condition for (c) that serves as a trigger for the intervener, the agent can consider alternatives in sense (a) without having his freedom overridden. The critic of *Global Revenge*, however, will almost certainly insist that a moral agent must have experience of considering alternate possibilities in sense (b), not just in sense (a).

I'm not at all sure that the objector isn't simply begging the question against the anti-PAPist by insisting on this requirement. But let me accept the challenge and suggest some examples, of varying value, designed to cast doubt on the idea that (G) undermines Jones's responsibility in *Global Revenge*.

The Supreme Agent. On some views, at least, God never considers alternatives before acting, though he is the supreme agent.¹⁶ If this is correct, then agency *per se* does not require a consideration of alternatives at some point in the agent's history, and the burden of proof is on the one who thinks that human agency in

particular requires such a consideration. This is all the more true if human agency is most fully agency insofar as it imitates divine agency.¹⁷

Mother Love. A mother who dies while rescuing her newborn child from a burning building is morally praiseworthy for so doing, even (perhaps especially) if she never considers doing otherwise, and the presence of a Frankfurt mechanism, programmed to make her decide to save the child if she considers not doing so, would do nothing to mitigate our praise. If the mother's moral responsibility in this case depends on there being some (earlier) occasion on which she considered doing A, considered not doing A, and chose between them, what would A have to be so that the requirement would be even remotely plausible—that is, that her praiseworthiness in this case could piggyback on *that*?

The Rehearsal. When a parent risks his or her life for a child, the parent's action seems *natural* in a way that might not require character formation. (Perhaps it's acting in the *contrary* manner that would require explanation in terms of earlier character-*deformation* episodes). But even where character development is relevant and indeed necessary, it isn't always necessary to achieve it through actual episodes of moral choice. Agents sometimes *rehearse* choices for later, in situations in which they are not in fact presented with the choice or even in a position to be presented with it. This happens when agents entertain "what-if" scenarios, reflect on the experiences of role-models who have been in similar situations, read edifying *contes moraux*, listen to sermons on the subject, or take an ethics course.

Tax Evasion. Of the various events that could serve as triggers for a counterfactual intervener, considering an alternative before undertaking to access it is arguably the one best qualified, in the greatest number and variety of cases, to fill the role of a genuinely necessary condition on doing otherwise. It is for this reason that it makes a good universal tripwire for the alternative-eliminator in *Global Revenge*. But there are also occasions on which it is plausible to suppose some further necessary condition for accessing a morally relevant alternative, beyond merely considering the alternative. For example, when a person has some initial inclination to do something, doing otherwise may require not only considering the alternative, but considering some reason for doing otherwise. Or when strongly tempted to do something, resisting the temptation may require not only considering the alternative and having some reason for choosing the alternative, but also attaching sufficient weight to the reason. Derk Pereboom's *Tax Evasion*, in which a man cheats on his taxes and might have refrained from cheating only if a moral reason had occurred to him with sufficient force, is a case of this sort. In such cases, the device may be programmed to intervene when and only when one of these further necessary conditions is satisfied, allowing Jones to consider both A and not-A without the device automatically interfering in his decision process.¹⁸

So much for the first response to the objection, which takes over the alternative-elimination strategy employed in *Revenge* and endeavors to apply it to earlier

episodes that are candidates for character-formation. The second response is that it's not clear that the appeal to earlier alternatives does the job even when such alternatives are available. The problem is that character-forming choices appear to explain the wrong thing when it comes to cases like *Revenge* (or, for that matter, *Assassin*). In *Sociopath*, for example, Jones is such that the kind and degree of humiliation he suffered at Smith's hands makes the decision to kill Smith psychologically irresistible. In this case Jones's moral responsibility may piggyback on earlier choices by which he became the kind of person who can't do other than murder Smith. (If there were no earlier choices of the right sort—e.g., if Jones's sociopathy were primarily the result of brain chemistry, or abuse he suffered as a child—this would mitigate Jones's responsibility, as PAP implies.) But *Revenge* is not *Sociopath*; the Jones of *Revenge*, absent the counterfactual intervener, is free to decide or not to decide to kill Smith. Whether Jones is the kind of person who would commit murder when humiliated, without giving it a second thought, depends on what he ends up doing, and so long as it is possible for him to do otherwise, it is possible for him to show by his actions that he is *not* that kind of person, despite earlier character-forming choices. *Adding the intervener changes nothing*. Jones now has no morally relevant alternative, but this is owing to the presence of the device, not to the formative power of some earlier choice whose satisfaction of PAP might apply, by extension, to Jones's choice in *Revenge*. The appeal to earlier character-forming choices does not successfully engage the challenge posed by buffer cases.

Conclusion

These seem to me to be the strongest grounds for thinking that *Revenge* harbors alternatives sufficient to satisfy PAP. I think these grounds are inadequate, for the reasons given above. Black the libertarian could participate in *Revenge* confident that a Jones who decided on his own to murder Smith could not have decided otherwise. If this is correct, then accessible alternatives are merely a normal *symptom* of morally responsible agency: they should not be identified, even in part, with the underlying condition of which they are symptomatic, any more than spots are to be identified with measles. And what is the underlying condition, in the case of moral responsibility? That's a very good question.¹⁹

Notes

1. Harry G. Frankfurt, "Alternate Possibilities and Moral Responsibility," *Journal of Philosophy* 66 (December 4, 1969), pp. 829–39.
2. *Ibid.*, p. 835.
3. *Ibid.*, pp. 836–7.
4. Locke's locked room example, from which blockage is extrapolated, may be found in *An Essay Concerning Human Understanding* Bk. II, ch. XXI, section 10.

5. I present the theological counterexample in “On Augustine’s Way Out,” *Faith and Philosophy* 16 (January 1999), pp. 3–26. The blockage strategy, together with further discussion of the divine foreknowledge scenario (understood as a species of blockage), may be found in my “Moral Responsibility and Unavoidable Action,” *Philosophical Studies* 97 (January (II) 2000), pp. 195–227, and “Freedom, Foreknowledge and Frankfurt,” *Moral Responsibility and Alternative Possibilities: Essays on the Importance of Alternative Possibilities*, eds. David Widerker and Michael McKenna (Aldershot, UK: Ashgate Pub. Ltd., 2003), pp. 159–83.
6. It might be objected that Jones could have done something about being *the kind of person* to whom such a thought would come unbidden in the specified circumstances; it’s therefore not true that he could not have avoided having this thought, for he *would* have avoided it if he had taken advantage of alternatives that were available to him in earlier character-forming episodes in his life. This objection can be circumvented with a slight modification to *Revenge*. Suppose the thought of killing Smith is planted in Jones’s mind by an external source—perhaps even by Black’s mechanism. So its coming to be there has nothing to do with Jones being one kind of person rather than another. But given that it’s there, Jones’s duty is now to accept or reject this idea, and nothing constrains him one way or the other. (A related objection appeals to earlier choices that formed Jones into the kind of person who would *accept* the idea of killing Jones, once the idea occurred to him. This objection cannot be handled by a simple revision in the terms of *Revenge*. I take up this objection later in the paper: see The Third Objection and Reply.)
7. I first presented a case like this one in “Moral Responsibility and Unavoidable Action,” where it took a backseat to my development of the blockage strategy. Derk Pereboom independently developed a similar case: his *Tax Evasion*, like *Revenge*, employs necessary rather than sufficient conditions as the mechanism’s trigger. See Pereboom’s *Living without Free Will* (Cambridge: Cambridge University Press, 2001), pp. 18–37.
8. What follows is largely taken from a section (“Three Objections Claiming the Presence of Alternatives”) of my “Moral Responsibility and Buffered Alternatives,” *Midwest Studies in Philosophy*, vol. 29: Free Will and Moral Responsibility, eds. Peter A. French & Howard K. Wettstein, guest editor John Martin Fischer (Boston & Oxford: Blackwell Pub., 2005), pp. 126–145. Readers interested in the first kind of objection, which I ignore in the present paper, will find it treated in another section of the *Midwest Studies* paper (“Three Objections Claiming the Absence of Moral Responsibility”).
9. I should make clear that I am not here proposing any necessary and/or sufficient conditions for the sort of moral relevance that would qualify an alternative to satisfy clause (ii) of PAP. John Davenport (“The Deliberative Relevance of Refraining from Deciding: A Response to McKenna and Pereboom”, *Acta Analytica* 21:4 (2006), 62–88) presents two cases—“One Lie Leads to Another” and “Pool Party”—which seem to me to raise interesting questions about whether any very simple account of the moral relevance in question can succeed. What I *am* claiming is that *Jones’s merely considering whether to murder Smith, in the context of Revenge (as specified)*, need not constitute an alternative satisfying PAP. Of course *Revenge* could be *modified* by introducing some of the psychological complexity to which Davenport draws our attention, and the result might well be more favorable to PAP—e.g., if Jones refrains from *considering* not killing Smith because he subconsciously fears that this might pave the way for his *deciding* not to kill Smith, we might agree that the availability of this unactualized alternative could then be relevant to his culpability for deciding to kill Smith. But this is a *different* case than the one I have described, and the existence of cases in which PAP is satisfied does nothing to rescue PAP from the Frankfurtian critique. PAP is in trouble if there is a single counterexample to it, and my claim is that *Revenge* (that very case, and not some neighboring case involving Sartrean “bad faith,” etc.) is just such a counterexample.

10. Daniel von Wachter made this objection when I presented the paper at Oxford in 2003.
11. Audiences at oral presentations of the paper have been more likely to raise this objection than any of the other five.
12. The basic objection conveyed in this paragraph was suggested to me (independently) by Robert Merrihew Adams and Gordon Pettit.
13. The idea that moral responsibility ultimately rests on earlier character-forming episodes which themselves satisfy PAP is most closely associated with Robert Kane; see especially his *The Significance of Free Will*.
14. I want to thank Gordon Pettit for raising this objection when I presented the first draft of this paper at the University of Notre Dame.
15. A Thomistic analysis of this strange individual might go like this. He has a *powerful intellect*, inasmuch as he is able to discern the good in any end as soon as he considers it; but he has a *weak will*, inasmuch as his appetite for the good is always moved by the first object proposed to it, before his intellect has a chance to consider alternative ends.
16. Descartes, for example, maintains in *Meditation IV* that God would have no need to *deliberate*.
17. This paragraph makes no mention of moral responsibility, which would apply to God (if it applies at all) in a very different sense than it applies to human beings. But it's not God's acting without considering the alternative(s) that makes his status with respect to moral responsibility different than that of human beings.
18. *Living without Free Will*, pp. 18–22.
19. Versions of this paper have been read at the University of Notre Dame, Westmont College, Oxford University, the University of California at Riverside, the University of Minnesota at Morris, and Talbot School of Theology, as well as at the 2000 Eastern Division meeting of the American Philosophical Association and the 2006 Bled Conference on Free Will. I want to thank the audiences at these venues as well as my commentators (Daniel von Wachter at Oxford, Michael McKenna at the APA).

Received: October 2006

Revised: November 2006

David P. Hunt
Department of Philosophy
Whittier College
Whittier, CA 90608
U.S.A.
Dhunt@whittier.edu