

## Frankfurt Examples, Derivative Responsibility, and the Timing Objection

Derk Pereboom, Cornell University

*Philosophical Issues* 22 (2012), pp. 298-315.

Penultimate draft

### Source views and Frankfurt examples<sup>1</sup>

Source views claim an agent's moral responsibility for an action is not to be explained primarily by the availability to her of alternative possibilities, but rather by the action's having a causal history in which she is the source of her action in a specific way. Agent-causal libertarianism is commonly conceived as an incompatibilist position in which an agent can be the source of her action in the way required for moral responsibility, and thus proponents of this view can be, and often are, source incompatibilists. However, one might also be a source incompatibilist and seriously doubt that we have the sort of free will required for moral responsibility, and this is the position I advocate.

Contemporary versions of source positions are typically motivated by Frankfurt examples (Frankfurt 1969). In examples of this sort, an agent considers performing some action, but a neuroscientist is concerned that she will not come through. So if she were to manifest an indication that she will not or might not perform the action, neuroscientist would

---

<sup>1</sup> Thanks to audiences at the University of Arizona, the University of Alberta, and the University of Rochester for valuable discussions, and thanks also to my free will seminar at Cornell in the Spring semester of 2011, and to David Widerker, Jonathan Vance, Seth Shabo, David Palmer, Dana Nelkin, Carlos Moya, Michael McKenna, Carl Ginet, Chris Franklin, Austin Duggan and Earl Conee for excellent comments.

intervene. But as things actually go, the neuroscientist remains idle, since the agent performs the action on her own. The idea is that even though the agent could not have avoided the action she performs, she is still intuitively morally responsible for this action.

Here is one of John Fischer's cases: Jones will decide to kill Smith only if Jones blushes beforehand. But his failure to blush by a certain time is what would trigger Black's intervention, which would cause him to kill Smith anyway. Supposing that Jones acts without intervention, we might well have the intuition that he is morally responsible for killing Smith, even though he could not have done otherwise than to kill him, and even though he could not even have formed an alternative intention.

Note that in this setup Jones does have an alternative possibility. She could have failed to blush. But Fischer argues that such a "flicker of freedom" is of no use to the Frankfurt-opponent, since it is not sufficiently *robust* to have a role in grounding the agent's moral responsibility. We might say that a robust alternative is one *relevant per se to explaining why an agent is morally responsible for an action*. Thus on Fischer's account, and on source views more generally, Frankfurt examples do not show that alternative possibilities aren't required for moral responsibility. Rather, what they indicate is that robust alternatives aren't required (Fischer 1994, 131-59).

A criterion for robustness

How should robustness be characterized more exactly? One intuition that underlies the proposal to ground moral responsibility in the accessibility of alternative possibilities is that to

be blameworthy for an action, an agent must have been able to do something that would have resulted in her being “off the hook” (Pereboom 2001: 1). In Carlos Moya’s phrasing, she must have an *exempting* alternative possibility, one that, should she avail herself of it, would exempt her from blame (Moya 2006: 67; cf. Otsuka 1998). A way of spelling this out more generally that I adopted earlier (2001) is that for an alternative possibility to be robust, it must be one that, had the agent taken it instead, would have precluded her from the responsibility she actually has for her action.

Robustness also has an epistemic component. Suppose that that the only way Joe could have avoided deciding to take an illegal deduction on his tax form -- a choice he does in fact make -- is by voluntarily taken a sip from his coffee cup, for unbeknownst to him, the coffee was poisoned, so a sip would have killed him immediately; or, for another version, it is laced with the drug that induces compliance with the tax code (Pereboom 2000, 2001; Moya 2006: 64 constructs an example of this form). In this situation, Joe could have behaved voluntarily so as to preclude the choice for which he was in fact blameworthy, as a result of which he would have been morally non-responsible for it. But whether he could have voluntarily taken the sip from the coffee cup is intuitively irrelevant to explaining why or whether he is morally responsible for his choice. What’s missing is that Joe has no inkling that taking the sip would render him blameless. This motivates the epistemic component of the robustness condition: if Joe were morally responsible because he has an alternative possibility in this situation, it must be that he in some sense understood that or how it was available to him.

Or suppose that that the only way Otto the Saxon could have avoided deciding to strike

his rival -- a choice he does in fact make -- is by instead voluntarily engaging in a fairly sophisticated meditative process that would have defused the anger that motivates him, whereupon he would not have decided to strike, and he would have been blameless. But he does not know anything about this meditative process – in fact, among the Saxons, no one is aware of it. Again, in this case if the agent was morally responsible for his decision because he has an alternative possibility, it must be that he in some way understands that it is available to him.

Thus here is an initial proposal (a substantial necessary condition) for what it is for an alternative possibility to be robust (Pereboom 2000; 2001: 26):

Robustness (A): For agent to have a robust alternative to her action A, that is, an alternative relevant per se to explaining why she is morally responsible for A, she must understand that instead she could have voluntarily done something as a result of which the specific sort of moral responsibility she actually has for A would have been precluded.

One might suggest that this for ‘the specific sort of moral responsibility she actually has for A would have been precluded’ we substitute, more simply, ‘she would have avoided doing A.’ But suppose that the agent could have done something just slightly different from A which for which she knows she would have been exactly as blameworthy as she would be for A. Such an alternative is not plausibly robust.

But here are a number of concerns for this formulation:

(a) Dana Nelkin poses the following challenge.<sup>2</sup> In a familiar example from Mark Twain, Huck Finn sincerely expresses the view that allowing Jim, the slave, to go free, is morally wrong, but nonetheless allows him to go free instead of returning him to his owner. We imagine that Huck has some cognitive and affective sensitivity to the moral rightness of letting Jim go free. Suppose that he instead, holding that moral psychology fixed, did return him to his owner, and that this is in fact morally wrong. Does Huck have a robust alternative possibility that can ground his moral responsibility – his allowing Jim to go free? My sense is that he might well, despite his not clearly understanding that letting Jim go free is morally right, and that letting him go free would have precluded him from the blameworthiness he actually merits. For, as Nelkin points out, we suppose that Huck has at least some cognitive sensitivity to the fact that letting Jim go free is morally right. As a result, Nelkin thinks that Robustness (A) is too strong, and I believe she is right. She suggests that understanding isn't required, but only some lower level of cognitive sensitivity to the fact that the alternative possibility would have precluded him from the responsibility she actually has.

(b) Jonathan Vance raises the following objection. In the coffee-cup example, is having a non-occurrent or even occurrent belief that taking a sip from the coffee cup *might* result in his being blameless enough for robustness?<sup>3</sup> It seems not. For, if asked, Joe might well agree that the probability of this connection is non-zero – he might admit, for instance, that it's at

---

<sup>2</sup> In correspondence; see also Pereboom (2009).

<sup>3</sup> In conversation. Kevin Timpe defended such a condition in the presentation of his paper "How Troublesome is Tracing" at the *Responsibility, Agency, and Persons* conference at the University of San Francisco in October, 2007; cf. Ginet 2000.

least .000001, and if he's taken a class in epistemology or probability, something like this might well be his response. But, intuitively, this is not sufficient to generate robustness. Should it be required for robustness that Joe understood that taking the sip of coffee would, with a probability of 1.0, result in his being blameless? This is clearly too strong, for it would intuitively be enough for robustness if he understood that the probability was, say, .95.<sup>4</sup> But the threshold probability, as one would expect, is difficult or impossible to determine.

(c) The agent's involuntarily becoming comatose instead of deciding to evade taxes would not count as a robust alternative. Thus it should be specified that the refraining or avoiding be voluntary (Hobbs 2012).<sup>5</sup>

(d) As Nelkin (2011) emphasizes, the requirement of an alternative that precludes actual responsibility is clearer and is more intuitive for immoral and blameworthy actions than for morally exemplary and praiseworthy actions.

On the assumption that in any situation taking the morally best option available renders an agent blameless, here is a proposal (again a substantial necessary condition – not a sufficient condition) that accommodates all of these concerns:

Robustness (B): For agent to have a robust alternative to her immoral action A, that is, an alternative relevant per se to explaining why she is blameworthy for A, it must be that

---

<sup>4</sup> Jonathan Vance made this point in conversation in September 2007, and Kevin Timpe argued for it in his presentation at the conference in San Francisco in October 2007.

<sup>5</sup> Hobbs makes this point in his Ph.D. dissertation. Recent discussions that emphasize the difference between positively doing otherwise and merely refraining from doing what one actually does include Steward (2009) and Alvarez (2009); cf. Pereboom (2001: 27).

(a) she instead could have voluntarily acted or refrained from acting as a result of which she would be blameless, and

(b) that for at least one such exempting acting or refraining, she is cognitively sensitive to its being available to her, with the result that she believes to some significant degree that had she voluntarily so acted or refrained she would be, or would likely be, blameless.

We can now characterize leeway views, whether compatibilist or incompatibilist, as those that affirm that for an agent to be blameworthy for an action, she must have available to her a robust alternative possibility, on my current best understanding, one that satisfies Robustness (B). Source views, by contrast, deny this. The leeway incompatibilist, whose position will arise in the discussion that follows, holds in addition that the reason causal determination precludes moral responsibility is that it rules out alternative possibilities altogether, but most importantly those of the robust sort.

#### Evading the dilemma defense

The most significant objection that has been raised against the earlier kinds of Frankfurt examples was initially suggested by Robert Kane and then systematically developed by David Widerker and Carl Ginet (Kane 1985: 51; 1996, 142-4, 191-2; Widerker 1995: 247-61; Ginet 1996). The general form of this Kane/Widerker/Ginet objection is this. For any Frankfurt example, if causal determinism is assumed to hold in the actual sequence that results in the action, the libertarian will not have and cannot be expected to have the intuition that the

agent is morally responsible. If, on the other hand, libertarian indeterminism in this actual sequence is presupposed, the scenario will not serve the Frankfurt-defender's purpose, for any such case will fall to a dilemma. In Frankfurt examples the actual situation will feature a prior sign that signals the fact that intervention is not required. If in the proposed case the prior sign causally determined the action, or if it were associated with some factor that did, the intervener's predictive ability could be explained. However, then the libertarian would not and could not be expected to have the intuition that the agent is morally responsible.<sup>6</sup> But if the relationship between the prior sign and the action was not causally deterministic in such ways, then the agent could have done otherwise despite the occurrence of the prior sign. Either way, an alternative-possibilities condition on moral responsibility emerges unscathed.

I have proposed a Frankfurt example that avoids this objection (Pereboom 2000; 2001: 18-19; 2003). Its distinguishing features are these: the cue for intervention is a necessary condition for the agent's availing herself of any robust alternative possibility (without the intervener's device in place), while the cue for intervention itself is not a robust alternative possibility, and the absence at any specific time of the cue for intervention in no sense causally determines the action the agent actually performs. Here is the example:

*Tax Evasion (2)*: Joe is considering claiming a tax deduction for the registration fee that he paid when he bought a house. He knows that claiming this deduction is illegal, but that he probably won't be caught, and that if he were, he could convincingly plead

---

<sup>6</sup> That is, unless the agent freely caused the occurrence of the prior sign, or was freely able to prevent its occurrence. Then the agent's responsibility for the action in question would plausibly be derivative of this prior free action or avoidance. The blush in Fischer's case is intended not to be subject to the agent's control in this way.

ignorance. Suppose he has a strong but not always overriding desire to advance his self-interest regardless of its cost to others and even if it involves illegal activity. In addition, the only way that in this situation he could fail to choose to evade taxes is for moral reasons, of which he is aware. He could not, for example, choose to evade taxes for no reason or simply on a whim. Moreover, it is causally necessary for his failing to choose to evade taxes in this situation that he attain a certain level of attentiveness to moral reasons. Joe can secure this level of attentiveness voluntarily. However, his attaining this level of attentiveness is not causally sufficient for his failing to choose to evade taxes. If he were to attain this level of attentiveness, he could, exercising his libertarian free will, either choose to evade taxes or refrain from so choosing (without the intervener's device in place). However, to ensure that he will choose to evade taxes, a neuroscientist has, unbeknownst to Joe, implanted a device in his brain, which, were it to sense the requisite level of attentiveness, would electronically stimulate the right neural centers so as to inevitably result in his making this choice. As it happens, Joe does not attain this level of attentiveness to his moral reasons, and he chooses to evade taxes on his own, while the device remains idle. (Pereboom 2000, 2001, 2003; David Hunt also suggested this 'necessary condition' strategy (Hunt 2000), and later develops a similar example (2005)).

In this situation, Joe is intuitively blameworthy for choosing to evade taxes despite the fact that he did not have a robust alternative.

Earlier types of cases, such as Fischer's blush example, also feature a necessary

condition for doing otherwise, i.e., the bluish not ever occurring. So the distinctive feature of *Tax Evasion* is not merely the presence of a necessary condition for doing otherwise. Rather, this case features a necessary condition for doing otherwise *the absence of which at any specific time will not causally determine the agent to perform the action*. This allows it to be the case that at no specific time is the agent causally determined to perform the action, which facilitates satisfaction of a key libertarian condition on moral responsibility. In *Tax Evasion (2)*, the necessary condition for not deciding to evade taxes, i.e. Joe's having the requisite level of attentiveness to the moral reasons, is the right sort, since its absence at any specific time does not causally determine his deciding to evade taxes. For at any particular time at which the level of attentiveness is absent, Joe could still make it occur at a later time, and thus he is not causally determined to decide to evade taxes by its absence at the previous time (although if there is a deadline, we'll need to think about what happens if he waits to decide until then, an issue that we'll encounter shortly).

The example does feature alternative possibilities that are available to the agent -- his achieving higher levels of attentiveness to moral reasons. Indeed, at this point one might object that given that the intervener's device is in place, by voluntarily achieving the specified higher level of attentiveness Joe would have voluntarily done something as a result of which he would have avoided the blameworthiness he actually incurs. Had he voluntarily achieved the requisite level of attentiveness, the intervention would have taken place, whereupon he would not have been blameworthy for deciding to evade taxes. But this alternative possibility is not robust. Joe is not believe to any significant degree, and, moreover, he has no appreciable

reason to believe, that by voluntarily achieving the requisite level of attentiveness he would not be (or likely not be) blameworthy. True, were he voluntarily to attain this attentiveness, the intervention would take place, and he would then not have been blameworthy at all. Still, Joe has no inkling, and no evidence, that the intervention would then take place, as a result of which he would not be blameworthy. In fact, he believes that achieving this level of attentiveness is compatible with his freely deciding to evade taxes anyway (which would be true if the intervener's device were not in place), and he has no reason to suspect otherwise. We can even specify that he believes that if he did achieve this level of attentiveness, he would still be very likely to decide to evade taxes. Nevertheless, intuitively Joe is blameworthy for actually so deciding.

Carlos Moya objects that in this situation Joe's blameworthiness should be determined by the following principle:

If, unbeknownst to her, an agent cannot do something A such that, if she did it, she would fulfill her duty and would be precluded from blame (and she knows that she would), then, in order to be so precluded, she should perform the next best action that reasonably was in her power to perform in order to fulfill her duty, where "the next best action" may be characterized, in general terms, as trying or attempting to A. (Moya 2011: 15)

The Frankfurt defender can agree that Joe is blameworthy for not becoming more attentive to the moral reasons, and that for this he does have a robust, exempting alternative possibility. But it's intuitive that Joe is also blameworthy for deciding to evade taxes, and for this, at least

prima facie, he has no robust alternative. What is the motivation for thinking that becoming more attentive to the moral reasons now becomes a robust alternative to deciding to evade taxes? Agreed: it is the next best action available to him. But it is not per se exempting for deciding to evade taxes, since, without the device in place, even if he did become more attentive he could decide to evade taxes, and, we might suppose, he even would be likely to so decide. And he is not at all aware that due to the device, becoming more attentive would preclude him from deciding to evade taxes, and would thus render him blameless for this decision. True, under the supposition that without the availability of a robust alternative to an action an agent cannot be blameworthy for performing it, it's natural to hold that if no robust alternative is available she would be blameless in her situation if she does the best she can. But the supposition is just what's being challenged by the Frankfurt defender. Moya says:

In normal circumstances, with no device lurking, the standards for exempting alternatives would have risen to deciding not to evade taxes and not evading them; merely attending to moral reasons would not have been enough; but since Joe could not have decided and acted that way, the standards lower to the next best action he could perform in order to fulfill his moral duties, which so becomes an exempting alternative (2011: 17)

The feature of the context that Moya claims to alter the standard from the usual one is the unavailability of what would under normal circumstances be an exempting alternative.

However, this claim is itself fueled by the robust alternative possibilities requirement, which is what's at issue. If Moya's proposal for altering the usual standard could be justified by

considerations independent of this requirement, then perhaps it should sway the Frankfurt defender. But the justification Moya sets out, while not implausible, is not independent of the requirement (2011: 13-15).

Stewart Goetz contends that in general, the absence of the relevant alternative possibility in a Frankfurt case can only be explained by the agent's being causally determined by factors beyond his control in the actual sequence, and that this is so in *Tax Evasion* as well (2005, 2009: 101-106). After all, that Joe decides to evade taxes is entailed by the elements of the set-up of the example, which are beyond Joe's control. However, there are two ways of thinking about causal determinism: (a) as involving the claim that events are *entailed by* propositions that describe preceding conditions and the laws of nature, and (b) as involving the claim that events are actually *produced* by such preceding conditions in accord with the laws of nature. A proposition describing the conditions in the set-up of a suitably constructed Frankfurt example entails that the action will come about, and so the action will then be determined by these conditions in sense (a). But in such a Frankfurt example not all of these conditions are actually operative in producing the action -- some remain merely counterfactual. And, in particular, when it is merely counterfactual intervention that ensures an action, it is not causally determined in sense (b). Terminologically, to my ear sense (b) is genuine *causal* determination by preceding conditions, while sense (a) involves determination by preceding conditions that need not amount to causal determination. In *Tax Evasion*, my sense is that the action is ensured so that the preceding conditions determine the action in sense (a) without

causally determining them (sense b).<sup>7</sup>

Mere derivative responsibility?

Biff decides to get drunk, understanding that when he is intoxicated he will no longer be able to avoid being abusive to his companions, and then when he is drunk he assaults one of them. In this example, Joe satisfies paradigmatic general conditions on moral responsibility at the time he decides to get drunk, but not when he is drunk and abusive. If he is morally responsible for the assault, it is only derivatively so – derivative, in particular, on his being non-derivatively morally responsible when he decides to get drunk, and his foreseeing that when he is drunk he is likely to be abusive (Ginet 2000).

Widerker (2006) proposes a challenge to the *Tax Evasion* argument in which he applies the distinction between non-derivative and derivative responsibility to the example, claiming that non-derivative moral responsibility is subject to a principle of alternative possibilities (PAP). Specifically, Widerker argues that in *Tax Evasion* Joe is only derivatively blameworthy for his decision to evade taxes:

[A] problem with Pereboom's example is that, in it, the agent is *derivatively* blameworthy for the decision he made, because he has not done his reasonable best (or has not made a reasonable effort) to avoid making it. He should have been more

---

<sup>7</sup> Joseph Campbell (2006) in effect points out that source incompatibilists who claim that non-responsibility transfers through determination in sense (a) cannot at the same time appeal to standard Frankfurt examples to rule out the leeway position, since in those cases features of the set-up determine the action in sense (a). According to my source incompatibilism, non-responsibility transfers through determination in sense (b), but not in sense (a).

attentive to the moral reasons than he in fact was – something he could have done.

And in that case, he would not be blameworthy for deciding to evade taxes, as then he would be forced by the neuroscientist so to decide. If this is correct, then Pereboom's example is a case of derivative culpability, and hence is irrelevant to PAP, which... concerns itself only with direct or nonderivative culpability. (Widerker 2006: 173; Ginet anticipates this objection in Ginet 1996)

First of all, there is a sense in which this is a dialectically unsatisfying response to a Frankfurt example, for the reason that it explicitly cites a leeway position in support of its verdict about Joe's responsibility. Joe is non-derivatively morally responsible only for not deciding to be more attentive to the moral reasons, for only relative to this decision is a robust alternative possibility available to him, and thus any moral responsibility he has for deciding to evade taxes must be derivative of this decision. One might thus apply the following alternative-possibilities schema to any example of this kind: the agent is non-derivatively morally responsible for acting or refraining at a particular time only if a robust alternative possibility relative to so acting or refraining are accessible to her at that time, and all other moral responsibility is derivative of such non-derivative responsibility. However, in the case of *Tax Evasion*, the concern is that if this schema is applied, the force of the example will not be engaged. More generally, this procedure stands to miss the force of a potential counterexample, and thus risks failing to engage a serious objection.

Still, this alternative-possibilities schema may get matters right. But in discerning whether it does, the drawbacks for imposing it on situations like Joe's must be assessed. One

might initially think that there won't be any, for one's intuitions about whether agents are morally responsible do not distinguish between non-derivative and derivative responsibility. Thus it may be intuitive that Joe is morally responsible for deciding to evade taxes, and not intuitive that he is non-derivatively as opposed to merely derivatively responsible. However, the paradigm for derivative responsibility is provided by the drunkenness just cited. In this example, while some of the uncontroversial general conditions on non-derivative moral responsibility fail to be satisfied when Biff is drunk, they are all met at the time he decides to get drunk. But Biff's situation differs significantly from Joe's. Biff has knowingly placed himself in a position in which some of the uncontroversial general conditions on non-derivative moral responsibility fail to be satisfied at relevant subsequent times. This is not true for Joe. One might think that if an agent cannot be sufficiently attentive to moral reasons, one such uncontroversial condition fails to be satisfied. But when at any given time Joe is not sufficiently attentive to moral reasons, he understands that it is open to him to become sufficiently more attentive at any later time. Thus Joe's situation diverges in a crucial respect from paradigm examples of derivative responsibility. This is an indication that the application of the alternative-possibilities schema to Joe in the way suggested by Widerker's objection is not appropriate (Pereboom 2009).

At the same time, Widerker points out that *Tax Evasion 2* would not be as prone to his derivative-responsibility concern if the necessary condition were changed so that it wasn't a morally impressive voluntary act.<sup>8</sup> He mentions imagining his stern father's face as such a condition. However, this might either have laudable moral import or not. Nelkin suggests Joe's

imagining being punished, for instance serving time in prison, as the necessary condition.<sup>9</sup> This condition can clearly lack laudable moral import. In accord with this suggestion, we can suppose that a necessary condition for Joe's refraining from deciding to evade taxes is his that he imagine, with a certain degree of vividness, being at least fairly severely punished for doing so. Revising the example accordingly would respond to Widerker's objection, since we are less likely to suppose that he is morally responsible for his decision to evade taxes only derivatively on his so imagining. Let me wait until introducing the next objection before presenting an example that features this sort of revision (for a version of *Tax Evasion* revised in this way, see Pereboom 2009). At the same time, I don't think this revision is necessary, and one might object to this change, as Moya does, by claiming that Joe is then not sufficiently responsive to moral reasons (Moya 2011: 19-20). I believe that this concern for the revised sort of example can be preempted by the right sort of set up, but nevertheless it is a concern worth bearing in mind.

Ginet's criticism.

In an objection to *Tax Evasion*, specifically to the (2000, 2001) version of this example, Carl Ginet, a defender of leeway incompatibilism, argues that at the precise time Joe makes the decision to evade taxes he might instead have been attending to the moral reasons instead, and that this alternative possibility is robust: "for had J taken it, he would at t1 have been refraining from a willing – to do B [decide to take the illegal deduction] right then – such that

---

<sup>8</sup> In conversation.

<sup>9</sup> In the same conversation.

by so refraining he would have avoided responsibility for doing B right then and would have been aware that he was avoiding responsibility for doing B right then (that being such an obvious implication of his not doing B right then, of which he of course would have been aware)" (Ginet 2002). David Palmer (2011) and Christopher Franklin (2011), also leeway incompatibilists, raise and develop this concern as well.

The force of this objection depends on a component of Ginet's (1996) general response to Frankfurt cases. He begins this discussion with the following schema for such examples:

Black sets up a mechanism that monitors Jones's actions and would cause Jones's doing B by  $t_3$  if Jones has not already done B by some deadline  $t_2$ . We must suppose that had this mechanism been triggered at  $t_2$  it would have causally necessitated Jones's doing B by  $t_3$  in such a way as to render Jones unable to avoid doing B by  $t_3$ , and that there was no time at which Jones knew or should have known about this mechanism. The mechanism is not triggered because Jones does B at  $t_1$ , before  $t_2$ . [Let 'B' stand for deciding to murder someone, say Smith.]

Ginet contends that Jones will not be responsible for doing *by  $t_3$*  "because, owing to the presence of Black's mechanism, Jones could not have avoided it, "but "he may be responsible for the obtaining of the temporally more specific state of affairs" -- doing B at  $t_1$  -- "which he could have avoided." One might fail, he thinks, to distinguish *Jones's doing B at  $t_1$*  and *Jones's doing B by  $t_3$*  because they are morally equivalent in the sense that Jones would be as blameworthy for the first as he would be for the second had Black's mechanism not been present. This is what allows our (mistaken) sense that Jones is blameworthy for making the

decision by t3 to be explained by his being blameworthy for making the decision precisely at t1. What explains Jones's blameworthiness is, in the correct analysis, his having an alternative to doing B at t1, and what he is really blameworthy for is doing B at t1.

I will now propose a two-part challenge to Ginet's objection. The first part disputes the claim that our sense that Jones is blameworthy for doing B by t3, or blameworthy at all, can be explained by his having an alternative possibility at t1. The second argues that a successful Frankfurt example can allow for this claim.

Let me begin by presenting a filled-out Frankfurt example that fits Ginet's schema, and also features the kind of change in the trigger for intervention suggested by Widerker and Nelkin:

*Tax Cut:* Jones can vote for or against a modest tax cut for those in his high-income group by pushing either the 'yes' or the 'no' button in the voting booth. Once he has entered the voting booth, he has exactly two minutes to vote, and a downward-to-zero ticking timer is prominently displayed. If he does not vote, he will have to pay a fine, substantial enough so that in his situation he is committed with certainty to voting (either for or against), and this is underlain by the fact that the prospect of the fine, together with background conditions, causally determines him to vote. Jones has concluded that voting for the tax cut is barely on balance morally wrong, since he believes it would not stimulate the economy appreciably, while adding wealth to the already wealthy without helping the less well off, despite how it has been advertised. He is receptive and reactive to these general sorts of moral reasons: he would vote

against a substantially larger tax cut for his income group on account of reasons of this sort, and has actually done so in the past. He spends some time in the voting booth rehearsing the relevant moral and self-interested reasons. But what would be required for him to decide to vote against the tax cut is for him to vividly imagine that his boss would find out, whereupon due to her political leanings she would punish him by not promoting him to a better position. In this situation it is causally necessary for his not deciding to vote for the tax cut, and to vote against it instead, that he vividly imagine her finding out and not being promoted, which can occur to him involuntarily or else voluntarily by his libertarian free will. Jones understands that imagining the punishment scenario will put him in a motivational position to vote against. But so imagining is not causally sufficient for his deciding to vote against the tax cut, for even then he could still, by his libertarian free will, either decide to vote for or against (without the intervener's device in place). However, a neuroscientist has, unbeknownst to him, implanted a device in his brain, which, were it to sense his vividly imagining the punishment scenario, would stimulate his brain so as to causally determine the decision to vote for the tax cut. Jones's imagination is not exercised in this way, and he decides to vote in favor while the device remains idle.

Suppose that  $t_3$  is the last moment that Jones, by his lights, can make a decision to vote prior to the expiration of the two-minute window, and that he actually decides to vote at  $t_1$ , a few moments before  $t_3$ . I claim that Jones is blameworthy for choosing to vote in favor of the tax

cut by  $t_3$  despite the fact that for this he has no robust alternative possibility.<sup>10</sup>

Alternative ways of setting up the example involve morally laudable necessary conditions for voting against, such as being more attentive to the moral reasons or meditating on what some morally exemplary person would do. My sense is that such conditions would also yield good cases.

One concern for *Tax Evasion 2* is that even with the intervener's device in place, Jones understands that at any specific time, into the indefinite future, he can avoid responsibility for a decision to evade taxes at that time by being attentive to the moral reasons instead, and this would count as a robust alternative possibility for each time into the indefinite future. The provision in *Tax Cut* that Jones is committed with certainty to voting by the last moment that by his lights he can make a decision to vote before the deadline, that is, by  $t_3$ , even though relative to this temporal interval he has no robust alternative possibility, corrects for this type of concern. With this new provision in place, it is still true that at any specific time prior to  $t_3$  he understands that by vividly imagining the punishment scenario he can avoid responsibility at that instant for deciding to vote in favor, which relative to those specific times would count as alternative possibilities. But it is not the case that relative to the entire interval up to and including  $t_3$  he has a robust alternative possibility. In particular, it is ruled out by Jones's commitment to deciding to vote by  $t_3$  that he not make a decision to vote at all during this interval by way of vividly imagining the punishment scenario.

In Ginet's view, because Jones has no alternative possibility to deciding to vote in favor

---

<sup>10</sup> I propose this example together with the second part of the reply to Ginet in Pereboom

by  $t_3$ , he will not be responsible for this particular state of affairs. But Jones is nevertheless responsible – blameworthy, in particular -- for making this decision at  $t_1$ , and for this he does have an alternative possibility. Specifically, at  $t_1$  he could have avoided making a decision right then. Again, in Ginet's view, (i) Jones's making his decision at  $t_1$  with the intervener's device in place and (ii) Jones's making this decision by  $t_3$  supposing the device is not in place are morally equivalent in the sense that the first is as blameworthy, and for the same reasons, as would be the second were the intervener's device not in place. This would allow our sense that Jones is blameworthy for making the decision by  $t_3$  to be explained by his being blameworthy for making the decision precisely at  $t_1$ . Will this work?

Suppose that factors beyond Adam's control causally determine him to decide to fire his rifle at his intended victim, and to make this decision precisely at some instant  $t_1$ . The leeway incompatibilist will maintain that Adam is not blameworthy for making this decision at this instant. But now imagine that factors beyond his control causally determine Adam so to decide, but in such a way that it is up to him at which instant during a two minute interval, beginning at  $t_0$  up to and including  $t_3$ , he makes the decision. More specifically, factors beyond his control causally determine him to have a desire to kill so strong that that as a result he will inevitably make this decision at some instant in this interval. He first has this desire just prior to  $t_0$ , and it would persist to  $t_3$  should he not decide before then, and this desire would not change in strength during the interval. This last instant,  $t_3$ , is the deadline because Adam believes that if he decides to fire any later, his intended victim will be beyond reach. Because he is cognitively sensitive to the strength of his desire, from the time just prior to  $t_0$  Adam

believes with certainty that he will decide by  $t_3$ . He understands that which of the instants he decides makes no moral difference, and so he is indifferent among them. Imagine he decides at  $t_1$ . Given incompatibilist sensibilities, leeway or source, Adam will not be *blameworthy* for deciding at  $t_1$ . Instead, he is in a morally neutral sense responsible for deciding at  $t_1$  rather than at another instant in the interval.

One might think that Adam's ability to refrain from deciding at  $t_1$  and  $t_2$  is strong evidence that he can continue to refrain up to and including the  $t_3$  deadline.<sup>11</sup> But, plausibly, Adam can refrain from choosing at  $t_1$  only because he knows he can decide later. By analogy, suppose my kids are very hungry, it's now  $t_0$ , I need to leave at  $t_3$ , and I committed with certainty to feeding them before I leave. Suppose I can refrain from feeding them at  $t_1$ . It's highly credible that this is only because I know that I can feed them a little later, and not because I can refrain from feeding them at all by  $t_3$ .

The reason the leeway incompatibilist must give for his not being blameworthy for deciding at  $t_1$  is that he has no (robust) alternative to making his decision by  $t_3$ . According to the leeway incompatibilist it has to be the unavailability of some alternative possibility that explains why Adam is not blameworthy for his decision. In this case, causal determination is what excludes Adam's blameworthiness, and the leeway incompatibilist maintains that in general, causal determination rules out blameworthiness because it precludes alternative possibilities. The only plausible candidate is the unavailability of an alternative to making the decision by  $t_3$ , and thus on the leeway incompatibilist view, this unavailability would have to be

---

<sup>11</sup> Thanks to Carl Ginet and Austin Duggan for raising this type of concern.

sufficient for Adam's not being blameworthy at t1.

We can now draw the following consequence for Jones's situation in *Tax Cut*. The leeway incompatibilist will not be able to defend the claim that Jones's deciding at t1 to vote in favor of the tax cut with the intervener's device in place is as blameworthy, and for the same reasons, as would be his deciding to vote in favor by t3 without the device in place. For with the device in place, the leeway incompatibilist cannot explain Jones's blameworthiness for making his decision at t1, only his responsibility in a neutral sense for making the decision at t1 rather than at some other available instant. Although, as in Adam's situation, Jones does have an alternative to deciding at t1 -- for example, continuing to deliberate at t1 -- this will be insufficient to explain Jones's blameworthiness for making his decision at t1. For Jones has no robust alternative to making his decision by t3, and as in Adam's situation, for the leeway incompatibilist this will be sufficient for Jones's not being blameworthy for making his decision at t1. But it's nonetheless our strong intuition that Jones is blameworthy for deciding to vote in favor at t1, for which the leeway incompatibilist now has no explanation.

Thus our sense that Jones is blameworthy for deciding to vote in favor of the tax cut cannot be accounted for by his being blameworthy for making his decision at t1 together with the alternative he has to doing so. On the leeway incompatibilist view, it now seems that Jones's blameworthiness for making his decision at t1 would have to depend on his having a robust alternative to making the decision by t3. But for this he has no robust alternative possibility. So it appears that *Tax Cut* issues a serious challenge to that position, and that this Frankfurt example can withstand Ginet's objection.

Waiting until the last instant?

Imagine that Jones decides to push the 'yes' button not at  $t_3$ , which is by his lights the last instant he can decide to push a voting button before the two minutes elapse. Suppose he understands that having any non-decision thought at  $t_3$  rules out his making a decision to vote, and he would then not vote. As specified, due to the penalty he is committed with certainty to voting. So now, at  $t_3$ , he can no longer vividly imagine the punishment scenario, and thus at  $t_3$  he has no robust alternative to the decision he actually makes. Yet it's intuitive that Jones's waiting until  $t_3$  does not render him blameless for his decision, and that for this he is as blameworthy as he would be if he made his decision at  $t_1$  instead.

In response, Ginet (1996), Franklin (2011) and Palmer (2011) would claim that at  $t_3$  Jones is only derivatively responsible for his decision. Franklin (2011) contends that any Frankfurt case of this sort falls to a dilemma: on the first horn, if the agent decides prior to the last moment, for example, if Jones decides at  $t_1$  in *Tax Cut*, he has a robust alternative; on the second horn, if he decides at the last moment, his responsibility is merely derivative. Ginet (1996) argues that all Frankfurt examples in general are subject to this sort of dilemma. In the case of *Tax Cut*, since the first horn can be countered, this argument fails. Still one might wonder whether on account of his not having an alternative at  $t_3$  Jones is really only derivatively responsible for his decision. Against this, it is clearly not the case that at  $t_3$  he is causally determined to decide to vote as he does by factors beyond his control. True, at this last instant Jones has put himself in a position in which he has no alternative but to vote in

favor, but he has knowingly done so by his libertarian free will. Perhaps the only reason to claim, even from the incompatibilist point of view, that he is merely derivatively responsible is that he lacks a robust alternative.

There would then be a sense in which the verdict of mere derivative responsibility would be dialectically suspect, again for the reason that it would explicitly cite a leeway position in support of its verdict about Jones's responsibility. This response would then appear to involve an imposition of the alternative-possibilities schema: an agent is non-derivatively blameworthy for doing A only if a robust alternative possibility relative to doing A is accessible to her at that time, one that explains her blameworthiness for doing A, and all other blameworthiness is derivative of such non-derivative blameworthiness. The concern is that if this AP-schema is applied, the force of the Frankfurt example will not be engaged. Such a claim for Jones's merely derivative responsibility for deciding at t3 is what a commitment to the requirement of robust alternative possibilities for moral responsibility would counsel, but this requirement is what's at issue. Against the application of this schema one might argue that Jones's deciding at t3 doesn't match the paradigm example of derivative responsibility, in which the agent is intuitively responsible for an action in a situation in which he fails to satisfy uncontroversial requirements for moral responsibility. Then the agent's responsibility for the action would clearly have to trace back to an earlier context in which he does meet such conditions. But unlike the agent who is drunk and assaults his companion, at t3 Jones's rationality in acting, for example, isn't deficient. The concern is that the appeal to derivative responsibility is motivated by the presupposition of the robust alternative possibilities

requirement itself, and that this undercuts its force.

At the same time, there is a potentially significant difference between Jones's deciding at t3 in *Tax Cut* and Widerker's target, Joe's deciding in *Tax Evasion 2*, as well as Jones's deciding at t1 in *Tax Cut*. If Jones decides to vote in favor of the tax cut at t1, he could have avoided this decision right then by continuing to deliberate instead, while the conditions are such that at t3 Jones can do nothing other than decide to vote in favor. One diagnosis of this difference is that, by contrast with t1, at t3 Jones is causally determined to vote in favor by his having decided to wait until t3 together with his awareness of the penalty for not voting. It might then be argued that on account of this causal determination even the source incompatibilist must agree that Jones is only derivatively responsible if he decides at t3. Still, since Jones would be non-derivatively blameworthy for deciding at t1 despite the fact that for this he has no robust alternative, the Frankfurt defender isn't forced to claim that Jones would be non-derivatively blameworthy for deciding to vote in favor at t3.

A final remark

For proponents of the leeway position, the accessibility of alternative possibilities is per se relevant explaining an agent's morally responsibility for an action. The *Tax Evasion* and *Tax Cut* examples yield reason to reject the leeway view. Note again that these examples nevertheless feature alternative possibilities, but not the robust sort. The source incompatibilist position is that for an agent to be morally responsible for an action, she must be the source of her action in the right way, and this requires that the action's actual causal

history not be deterministic. More precisely, the action cannot be causally determined by factors beyond the agent's control, such as the remote past together with the laws of nature. It might well be that the agent's having alternative possibilities – but not necessarily of the robust sort -- are entailed by this constraint (Pereboom 2001: 37, 2003: 197; cf. Della Rocca 1998).

## References

- Alvarez, A. (2009) "Actions, Thought-experiments, and the 'Principle of Alternate Possibilities,'" *Australian Journal of Philosophy* 87, pp. 61-81.
- Campbell, J. (2006). "Farewell to Direct Source Incompatibilism," *Acta Analytica* 21, pp. 36-49.
- Della Rocca, M. (1998). "Frankfurt, Fischer and Flickers," *Noûs* 32, pp. 99-105.
- Fischer, J. M. (1994). *The Metaphysics of Free Will*, Oxford: Blackwell.
- Frankfurt, H. (1969). "Alternate Possibilities and Moral Responsibility," *Journal of Philosophy* 66, pp. 829-39.
- Franklin, C. (2011). "Neo-Frankfurtians and Buffer Cases: The New Challenge to the Principle of Alternative Possibilities," *Philosophical Studies* 152, pp. 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, pp. 403-17.
- Ginet, C. (2000). "The Epistemic Requirements for Moral Responsibility," *Philosophical Perspectives* 14, pp. 266-77.
- Ginet, C. (2002). "Review of *Living Without Free Will*," *Journal of Ethics* 6, pp. 305-9.
- Goetz, S. (2005). "Frankfurt-Style Arguments and Begging the Question," *Midwest Studies in Philosophy* 29, pp. 83-105.
- Goetz, S. (2009). *Freedom, Teleology, and Evil*, Continuum.
- Hobbs, J. (2012). Dissertation, Cornell University, forthcoming.
- Hunt, D. (2000). "Moral Responsibility and Unavoidable Action," *Philosophical Studies* 97, pp. 195-227

- Hunt, D. (2005). Moral Responsibility and Buffered Alternatives, *Midwest Studies in Philosophy* 29, pp. 126-45.
- Moya, C. (2006). *Moral Responsibility: The Ways of Skepticism*, Oxford, Oxford University Press. pp. 195-210.
- Moya, C. (2011). "On the Very Idea of a Robust Alternative," *Critica* 43, pp. 3-26.
- Otsuka, M. (1998). "Incompatibilism and the Avoidability of Blame," *Ethics* 108, pp. 685-701.
- Nelkin, D. (2011). *Making Sense of Freedom and Responsibility*, Oxford: Oxford University Press.
- Palmer, D. (2011). "Pereboom on the Frankfurt Cases," *Philosophical Studies*, forthcoming.
- Pereboom, D. (2000). "Alternative Possibilities and Causal Histories," *Philosophical Perspectives* 14, pp. 119-37
- Pereboom, D. (2001). *Living without Free Will*, Cambridge: Cambridge University Press.
- Pereboom, D. (2003). "Source Incompatibilism and Alternative Possibilities," in *Freedom, Responsibility, and Agency: Essays on the Importance of Alternative Possibilities*, M. McKenna and D. Widerker, eds., Aldershot, U.K.: Ashgate Press, 2003, pp. 185-99.
- Pereboom, D. (2009). "Further Thoughts about a Frankfurt-Style Argument," *Philosophical Explorations* 12 (2010), pp. 109-18
- Steward, H. (2008). "Fairness, Agency, and the Flicker of Freedom," *Noûs* 43, pp. 64-93.
- Widerker, D. (1995). "Libertarianism and Frankfurt's Attack on the Principle of Alternative Possibilities," *The Philosophical Review* 104, pp. 247-61.
- Widerker, D. (2006). "Libertarianism and the Philosophical Significance of Frankfurt Scenarios," *Journal of Philosophy* 103, pp. 163-87.