



Frankfurt-Style Counterexamples and the Importance of Alternative Possibilities

Nadine Elzein¹

Received: 8 March 2016 / Accepted: 22 July 2016 / Published online: 21 September 2016
© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract Proponents of modern Frankfurt-Style Counterexamples (FSCs) generally accept that we cannot construct successful FSCs in which there are no alternative possibilities present. But they maintain that we can construct successful FSCs in which there are no *morally significant* alternatives present and that such examples succeed in breaking any conceptual link between alternative possibilities and free will. I argue that it is not possible to construct an FSC that succeeds even in this weaker sense. In cases where any alternatives are *clearly* insignificant, it does not appear at all obvious that the agent can be held responsible. Present popular FSCs include alternatives that are ambiguous in their significance, and when the examples are sharpened to remove this ambiguity, they lose their force. Moreover, the proponent of such examples faces a problem: We can easily construct scenarios in which any alternatives are *obviously* insignificant, and in such scenarios, we are not intuitively inclined to suppose the agent is responsible. The proponent of new FSCs must therefore distinguish any alternatives she includes from the sorts included in these scenarios. The difference must now be such that (a) this helps to make it seem intuitively likely that the agent is responsible where the agent otherwise would not appear responsible, and (b) these alternatives are irrelevant to any judgment about whether the agent is responsible. I maintain that it is impossible to achieve both of these goals at once.

Keywords Moral responsibility · Frankfurt-style counterexamples · Leeway incompatibilism · Source-incompatibilism · Alternative possibilities

✉ Nadine Elzein
nadine.elzein@kcl.ac.uk

¹ Department of Philosophy, King's College London, The Strand London, WC2R 2LS UK

1 Introduction

The idea that alternative possibilities are important for moral responsibility has been steadily in decline since the publication of Harry Frankfurt's famous counterexample to the Principle of Alternate Possibilities, or "PAP" (Frankfurt 1969). But the nature of the dispute surrounding Frankfurt's argument has shifted in recent years.

Frankfurt's original example was typically seen as a straightforward counterexample to any principle which says that alternative possibilities are a necessary condition of moral responsibility. Frankfurt's stated target, PAP, is an example of one such principle, but there are many others.¹ Let's call all such alternative-possibility-invoking principles API principles.

The important class of principles might naturally seem modal in nature: They tell us that there are no possible worlds in which it is true both that some agent is morally responsible, and that this agent has no alternative possibilities available to him. If we can imagine a scenario in which an agent would count as responsible despite a lack of alternative possibilities, then this would provide an adequate counterexample to API principles, so understood. Understood like this, the force of Frankfurt's example would depend crucially on whether we can convincingly construct a scenario in which there are no alternative possibilities present *at all*, but in which the agent seems obviously responsible. This task has proved notoriously difficult.

In recent years, it has been broadly accepted that it is impossible to formulate a Frankfurt-Style Counterexample (FSC) in which there are no alternative possibilities present *at all*, without presupposing a deterministic setting, which would arguably make it question-begging to insist that the agent can nonetheless be held responsible.²

However, a wave of recent defenders of Frankfurt's argument ("neo-Frankfurtians", as Franklin 2011 has termed them) insist that FSCs could be effective in discrediting all API principles, even if it is impossible to formulate an example in which an agent appears responsible, and in which there are no alternative possibilities present.³

An FSC may still be able to undermine the claim that alternative possibilities *per se* have any essential role in actually *supporting* the agent's moral responsibility. This could be done by providing an example in which any alternatives that *are* present can be shown to be irrelevant to the agent's moral responsibility, and in which the agent appears to be responsible nonetheless. Such alternatives may be regarded as irrelevant because they lie outside of the agent's control or otherwise lack moral importance. In these cases, the alternatives can plausibly be described as "insignificant" or "non-robust".

¹Some objections to Frankfurt's original argument specifically draw on a defence of alternative principles to PAP. For notable strategies of this sort, see van Inwagen (1978) and Otsuka (1998).

²Notable versions of this objection have been put forward by Kane (1985, p. 51; 1996, pp. 142-4, pp. 191-2), Ginet (1996) and Widerker (1995). For a recent defence, see Goetz (2005, 2008, pp. 75-122).

³Note that some theorists attack the other horn of the dilemma, arguing that it is not question-begging to presuppose a deterministic setting. This line of argument has been defended by Fischer (1999, 2006, 2010), and by Haji and McKenna (2004, 2006). I will not address this argument here. My focus is solely on examples which aim to avoid any assumption of determinism, and to instead get around the dilemma by presenting a case in which any alternatives present are non-robust.

Examples of this sort may certainly help to undermine the claim that alternatives have any essential role in *grounding* responsibility, even if they do *not* establish that there are possible worlds in which the two come apart. This could either be a first step in an argument for compatibilism, or else push us towards a different form of incompatibilism—one which states that if determinism conflicts with free will, it is *not* because agents lack alternative possibilities in deterministic contexts. Instead, the absence of alternatives may be an indicator that some other condition of freedom is lacking, e.g., “ultimate sourcehood” or “origination”.

Further to opening up new resources for Frankfurt-inspired compatibilism then, this development has also divided incompatibilists into two camps: “Source incompatibilists” argue that alternative possibilities are not important *per se*, but merely as indicators of agents’ status as the ultimate source of their own decisions.⁴ “Leeway incompatibilists” defend the traditional view that alternative possibilities are central to moral responsibility either in their own right, or because they are in some way essentially bound up with sourcehood, so as to make the two sorts of concern inseparable.⁵

Following Leon and Tognazzini (2010), we might say that a “weakly successful” FSC is one in which an agent is responsible despite the fact that any alternatives available are non-robust. This contrasts with a “strongly successful” FSC, in which an agent is responsible despite the fact that there are no alternative possibilities present *at all*. Neo-Frankfurtians claim that FSCs can refute API principles so long as they are *weakly* successful. The idea that we should require a strongly successful example appears to be naive: it requires us to suppose that a robust notion of moral responsibility can be pinned on trivial or irrelevant alternatives.

I believe, however, that there are no examples, and probably *could* be no examples, that succeed, even in the *weak* sense; that there are no cases in which it is obvious that an agent is morally responsible, but in which there are only insignificant alternatives present. I need not then, endorse the naive view that insignificant alternatives can ground a robust notion of moral responsibility.

It is very easy to construct examples in which any alternatives left open are *obviously* irrelevant to an agent’s responsibility. But in these examples, it is far from obvious that the agent can be held responsible. In fact, popular present FSCs include alternatives that are not *obviously* irrelevant at all, but which are at best ambiguous or obscure in their relevance. I will argue that these FSCs depend entirely on this obscurity for their force. This hypothesis is easily tested: We can test it simply by sharpening the examples so as to remove the ambiguity, and then seeing if those examples remain convincing.

I wish to compare present FSCs with what I will call “simple cases”. Simple cases are scenarios in which it is unambiguously clear and obvious that there are only irrelevant alternatives present.

⁴Proponents of his view include Stump (1999a, b, 2003), Pereboom (2000, 2001, 2003, 2009), Zagzebski (2000, 2010), Timpe (2007, 2008), Shabo (2010), and Widerker (2006, 2009).

⁵Recent defender’s include Kane (2000, 2002, 2004, 2007a, b) Moya (2006, 2007, 2011), Alvarez (2009), Ekstrom (2003), and Steward (2009). See also Goetz (2008, pp.75–122).

There are two main grounds on which an alternative may be deemed irrelevant (both of which are commonly utilised by neo-Frankfurtians). Firstly, an agent's options may be morally undifferentiated, such that she will be equally blameworthy or praiseworthy regardless of which she chooses. Secondly, an agent may be unable to access an alternative through her own efforts, which will similarly render the alternative morally irrelevant. It is very easy to think up cases in which it is absolutely clear that any alternatives present are irrelevant in either of these respects. Consider first the case of agents choosing between morally undifferentiated options.

If an alternative is going to be relevant to an agent's moral responsibility, then there must be some difference in moral value between what the agent does and what the agent might have done, and the agent must be aware of this difference. If the moral merits of the agent are exactly the same whichever way she chooses, or if the agent does not know that there is a moral difference between the alternatives and could not be expected to know this (which amounts to the same thing), then this alternative will be morally insignificant. An easy source of simple scenarios, in which there is no moral difference between two options, are those in which the agent could not possibly *know* that whether she later misbehaves will depend on which of these alternatives she chooses. Let's consider a clear-cut case of this sort.

Simple Case 1:

Suppose that Brenda lives in a universe in which her actions are entirely determined by factors outside of her control, save for just one event: At a particular time t , she can take a sip from a cup of coffee, or else refrain from taking a sip. This is the only genuinely chancy event that ever occurs in her universe, and hence, this is the only libertarian free choice she ever makes in her life. If Brenda does not take a sip from her coffee at t , then at a later time $t1$, she will slash her boss's car tyres in a fit of spite. However, unbeknownst to Brenda, her coffee is spiked with heavy sedatives, and so were she to take a sip, her later annoyance would never escalate into a fit of spite, and she would never develop the motivation to slash her boss's tyres. There is absolutely no other way that Brenda could possibly avoid choosing to slash her boss's tyres at $t1$, except by taking a sip of coffee at t . Brenda does not know her coffee is spiked and would have this effect, and there is no way she could possibly be expected to know this, and so she does not realise that there is any important moral difference between the scenario in which she takes a sip from the coffee and the scenario in which she does not. Unaware as she is, she does not feel inclined to drink coffee today, and so does not take a sip. As a result, she is then unable to avoid slashing her boss's tyres in a fit of spite.

In this example, I have gone to great pains to make absolutely clear (a) that there is only one alternative available to Brenda and (b) that this alternative is morally irrelevant. And in this example, it does not seem at all clear that we can assert (without at least some controversy), that Brenda is responsible. Perhaps we ought to expect some disagreement in intuitions here. But, at the very least, it seems that anyone who is strongly inclined to think that it would be problematic to insist that she is responsible in a completely determined world, is likely to find it *similarly* problematic to insist that this world fares any better. Since the coffee-sipping alternative is utterly

irrelevant to her moral worth, it seems that adding this alternative, and this alone, adds nothing to her moral responsibility over what would be present in a deterministic world.

Of course, to turn this into an FSC, we need only stipulate that if she takes a sip from her coffee some intervener will force her to slash her boss's tyres anyway. But this additional step is unnecessary for our purposes: All that needs to be shown is that if this is her *only* alternative, and it is made obvious that it really *is* irrelevant, then it will be far from obvious that we can hold her responsible, irrespective of the presence of the potential intervener.

Another way in which alternatives could be irrelevant to the agent's responsibility is if it is not within the agent's power to do anything that would determine which of those alternatives comes about. As Fischer puts it, "there is a crucial difference between the *ability* to do otherwise and the *mere possibility* of something different happening" (Fischer 2003, p. 196). If the only way, an agent could avoid responsibility is by unwillingly fainting at the crucial moment or being struck by lightning, then this sort of alternative will not be significant as far as the agent's responsibility goes. We may call alternatives of this sort "inaccessible" to the agent.

It is very easy to think up examples in which there are obviously only inaccessible alternatives present. Here is one such example:

Simple Case 2:

Let's suppose that Brenda lives in a completely determined world, save for one chancy event: At t , Brenda could be struck down by a lightning bolt from the Gods, which is the only chancy occurrence in an otherwise entirely deterministic universe. Brenda will definitely slash her boss's tyres in a fit of spite at $t1$ unless the Gods strike Brenda with lightning at t (which will seriously injure her, ensuring that she never even begins deliberating about it). This is the only way that Brenda could possibly avoid deciding to slash her boss's tyres at $t1$.

Given that Brenda's choices and actions are entirely determined, save for the possibility of something happening to her which is entirely outside of her control, it does not look at all obvious that Brenda is responsible for slashing her boss's tyres. Again, anyone who is intuitively inclined to worry that Brenda is not responsible for her actions in a deterministic universe, without the lightning-bolt possibility, will be equally inclined to doubt that she is responsible in *this* universe, where this feature is the only difference, and its irrelevance to her moral agency is absolutely clear.

Neo-Frankfurtians must try to formulate examples which are like these simple cases, in that any alternatives present are irrelevant to the agent's responsibility, but which differ from these cases, in that they are able to reliably produce the intuition that the agent is nonetheless responsible. Some try to show that (like in Simple Case 1) an agent may have only morally irrelevant alternatives available, but may (unlike in Simple Case 1) nonetheless appear to be responsible. In this category, we have examples such as those provided by Pereboom (2009, 2003, 2007a, b), Hunt (2005), and McKenna (2003). I will focus on McKenna and Pereboom's examples. Since Hunt's example is very similar in structure to Pereboom's, it should be obvious how my arguments might be generalised.

Another strategy is to try to show that there can be scenarios in which (like in Simple Case 2) any alternatives are inaccessible, but where (unlike in Simple Case 2) the agent might nonetheless appear responsible. In this category, we have examples such as those provided by Hunt (2000), Stump (1999a, b, 2003), Widerker (2006, 2009), Fischer (1999, 2003, 2006, 2007) and Robb and Mele (1998, 2003). I do not have the space to discuss all of these formulations or others in the literature. I will instead focus just on the sort of example put forward by Fischer and Stump. But I hope to show that my arguments can be generalised to the others.

My strategy when it comes to dealing with these three examples will follow a particular three-step pattern, which should soon become familiar:

Step one:

I will try to identify the important feature that differentiates the FSC under evaluation from the simple case, enabling it to produce the intuition that the agent is responsible, where we do not have such an intuition in the simple case. For each FSC, I will argue that the relevant distinguishing feature relates directly to the sorts of alternative in play: The simple case features alternatives that are obviously irrelevant, whereas the relevance of the alternatives in the FSC is unclear.

Step two:

In each case, I consider arguments offered in support of the conclusion that we should regard the alternatives as irrelevant. I aim to show that these arguments are either weak or else *can* be made convincing, but that doing so damages any intuition that the agent is responsible.

Step three:

Finally, I will show that these examples rely crucially on preserving an ambiguity about the relevance of the alternatives in play. We can demonstrate this by sharpening the example so as to remove the ambiguity: We can either make it absolutely clear that the alternatives *are* relevant (i.e. that they differ from one another in moral value, that the agent is aware of this difference and that the alternative is accessible to the agent), or we can make it absolutely clear that the alternatives are *not* relevant (i.e. that the alternatives are inaccessible or that the agent could not be aware of any moral difference between the alternatives). When we do this, we find that agents seem obviously responsible when the alternatives are made obviously relevant, and that it looks very doubtful that they are responsible when the alternatives are made obviously irrelevant.

This should help bring into focus a more fundamental difficulty. Given that the relevant distinguishing feature in these cases relates directly to the nature of the alternatives involved, we will only have a successful example if we can meet the following requirements:

1. Our FSC is distinguished from the simple case by virtue of having alternatives with some feature (or set of features) that enables it to reliably produce the judgement that the agent is responsible.
2. These features are completely irrelevant to the agent's responsibility.

3. These features are able to produce this judgement without the help of confusion or unclarity on the question of their relevance.

I contend that these requirements cannot possibly all be fulfilled at once; that is, that no feature or set of features can, on its own, *reliably produce* a judgement, and do so for clear and principled reasons that do not rest on any ambiguity or confusion, whilst at the same time being *utterly irrelevant* to that judgement.

Of course, if the relevant distinguishing feature which accounts for the disparity in judgements between these cases were completely independent of the nature of the alternatives involved, the examples could avoid this conceptual problem. But I hope to make clear that there are, at least in these examples, no plausible candidates for such a feature.

Some theorists might object to my evaluating FSCs in this way. They may argue that whether an example evokes the correct judgement is beside the point: What matters is whether we can present a successful *argument* to the conclusion that the agent ought to be regarded as responsible. But this objection is misguided. Suppose that in our FSC, it is clear that there are only insignificant alternatives available, but it seems doubtful that the agent is morally responsible. But suppose that we can offer an argument to the conclusion that we *ought* to hold the agent responsible, even if this is counterintuitive. We would now have an *independent* argument to the conclusion that agents ought to be regarded as morally responsible despite an absence of robust alternatives. Since the success of our FSC would be dependent on the success of this independent argument, this would render the FSC itself redundant. My argument is precisely that FSCs do not help to provide any *basis* for rejecting API principles.

The remainder of this paper will be divided into four sections. In the first, I examine FSCs that are alleged to include only morally undifferentiated alternatives. In the second, I examine FSCs that are alleged to include only only inaccessible alternatives. I will argue that all of these formulations face a serious problem, and that this objection can be generalised to apply to other FSCs. In the third section, I will discuss some possible objections to my argument. Finally, I will conclude by offering some remarks about the overall point of my argument and its implications.

2 FSCs and Morally Irrelevant Alternatives

2.1 McKenna's FSC

McKenna offers the following example:

Tax Deadline:

Suppose Betty deliberates over the option of cheating on her taxes. She considers the options of following or violating the tax law. She might further deliberate about the advantages of one or the other, such as what she might buy with the extra money or whether she would have to do jail time if she were caught. But, granting Betty libertarian freedom, surely there is a range of alternative courses of action to deliberately cheating or deliberately complying with the law that are simply not relevant to Betty's deliberative circumstances.

Suppose that there is some time deadline and that Betty must make this decision quite soon (it is early afternoon and Betty must have her tax returns at the post office by 5:00 in the afternoon). Betty might simply stop all this nonsense, cease deliberating (that is, *not* decide) and head for the gym, or instead, roast a chicken. But these alternative are not morally significant given Betty's deliberative perspective. As it turns out, Betty does cheat on her taxes. (McKenna 2003, pp. 206–207).

Suppose that Betty can escape cheating her taxes *only* by getting distracted and failing to come to a decision on the matter. Irrespective of any counterfactual intervener, we are supposed to find it immediately plausible that Betty is responsible, even though her *only* alternative is non-robust. So let us go through the three-step strategy outlined at the start.

Step 1:

Are there are any obvious differences between this example and Simple Case 1, which might account for the fact that Betty seems to be responsible in this example, where Brenda, in our simple case, does not?

Firstly, to recap, in Simple Case 1, Brenda had the following alternative: Either take sip from her coffee, or refrain from taking a sip from her coffee. Brenda had no reason to suspect that taking a sip from her coffee would be her only possible way of avoiding some later wrongdoing. In this case, Betty can either continue deliberating about whether to cheat her taxes, or else she can choose to roast a chicken or go to the gym. It is unclear whether she is supposed to realise that the only way in which she can possibly avoid later moral wrongdoing is by roasting a chicken or going to the gym. She is, presumably, perfectly aware that the latter course of action is morally innocent in comparison to actually cheating her taxes, and she is aware, presumably that she definitely will not cheat her taxes if she does any of these innocent things instead. We do not whether she realises that cheating her taxes is the inevitable result of continuing to deliberate.

One obvious difference between these cases immediately jumps out. It looks much less obvious in this example that Betty's alternatives are irrelevant. This looks less obvious because Betty appears to have relatively innocent alternatives to cheating her taxes, which she could avail herself of but chooses not to, and presumably Betty is aware of this moral difference. If Betty knows that roasting a chicken or going to the gym are morally innocent alternatives to cheating her taxes, she appears to have a rather robust moral choice to make.⁶ So, it looks on the face of it as if Betty's alternatives are in fact morally significant. In contrast, Brenda could not possibly have grasped the moral significance of her decision about whether to drink her coffee. A choice between possibly illegally cheating one's taxes or else going to the gym (and definitely *not* cheating one's taxes) is obviously more morally robust than a choice between sipping one's coffee or not sipping one's coffee.

⁶I will assume that roasting a chicken is morally better than cheating your taxes, although I note that from the chicken's perspective, this is hardly obvious!

McKenna urges us to resist this conclusion, arguing that the options are not in fact significant as they may seem. So let's move on to step 2 and examine this argument.

Step 2:

McKenna emphasises the fact that in order to avail herself of these alternatives, Betty must fail to even arrive at a decision about whether to cheat her taxes. In this scenario, the alternatives do not even allow her to understand or compare the competing moral worth of the different options. She can avoid cheating her taxes only by failing to even reach a decision, having become distracted by some trivial alternative. He argues that in this context, the alternatives are not robust, since they are not "relevant to *competent moral and deliberative agency*" (McKenna 2003, p. 207). But on this point, he says something rather puzzling:

Imagine how it would appear if things were otherwise. There is Betty in court, under interrogation. The prosecuting attorney says to Betty, 'Betty, did you at least consider not cheating on your taxes before sending them in?' Betty replies, 'Well, no, but I was thinking that I might roast a chicken.' This would not help to reinforce Betty's competency as a morally and legally responsible agent (McKenna 2003, p. 207).

What seems puzzling here is that we are supposed to have the intuition that Betty is nonetheless a responsible agent if we take this argument seriously. It seems odd to hold her responsible when the *only way she can possibly avoid* choosing to cheat her taxes is by roasting a chicken. This assumption is particularly strange given that McKenna himself has already stressed that if these were the only alternatives she *even considered*, she would not appear to be a morally responsible agent at all. Let alone if, as has actually been stipulated, Betty is *incapable* of considering alternatives that have any more relevance.

To the extent that we find the argument for the irrelevance of the alternatives convincing then, the claim that Betty is nonetheless responsible becomes correspondingly less compelling. This might lead us to suspect that McKenna's case only fares better than the simple case in terms of producing the intuition that the agent is responsible because it features alternatives of ambiguous moral relevance, where the simple case only features clearly irrelevant alternatives. So finally, let's test this hypothesis by sharpening the example and seeing what effect this has.

Step 3:

First then, we can make it much more obvious that this is Betty's only alternative and that it really is deliberatively insignificant. Suppose Betty says to the judge 'I'm sorry your honour, but it turns out that it was absolutely impossible for me to even *consider* not cheating my taxes. I have excellent evidence from neuroscientists evaluating my case that I could not possibly have done other than choose to cheat my taxes or else stop deliberating and roast a chicken, and at the time, I had no reason to suspect that roasting a chicken was my only way out of inevitably choosing to cheat my taxes. That is why I did not take this into account as a morally important option'. Well, now this looks like rather a *good* excuse. If her only way out was *genuinely* to do something trivial, which she did not know would have any bearing on her decision

(a strange condition, granted, but the one that must be stipulated if the example is to work at all), then it does not seem fair to hold her responsible.

Now the example is, effectively, just like the simple case: Her later wrongdoing rests purely on a decision about whether or not she roasts a chicken, and she has no reason at all to suspect that this decision is morally relevant. In this respect, it becomes much like Brenda's decision about whether to take a sip from her coffee. The only difference is that Betty's choice must disrupt a process of deliberation that is already underway. But so long as (a) Betty cannot reasonably be expected to know that this process will inevitably culminate in wrongdoing unless she now embarks on roasting a chicken and (b) has no other possible means of avoiding the wrongdoing, it seems similarly clear that she can hardly be held responsible.

So what if we instead sharpen it in the opposite direction? We could instead stipulate that it is totally clear to Betty that her options are either to cheat her taxes or to roast a chicken. Suppose someone puts a gun to her head and tells her that she must do one or the other of these two things, but it is entirely up to her which. And finally, suppose that roasting a chicken is morally better than cheating her taxes and she is fully aware of this. This certainly brings back the intuition that Betty is morally responsible for cheating her taxes as opposed to roasting a chicken. But that is only because she had a morally relevant alternative. Even if we stipulate that should she start roasting a chicken, an intervener will manipulate her brain and make her evade her taxes, she will still plausibly be responsible for deliberately choosing the morally worse of two options in the case where there is no need for intervention.

We can sharpen the example in both directions then, and in neither case do we have a convincing FSC. Here, we should ask a simple question: If we do not suppose that the agent is responsible in cases where the alternative is clearly irrelevant, such as Simple Case 1, why should changing the nature of the irrelevant alternative improve matters? If the feature that distinguishes this case from the simple one also enables it to elicit the judgement that the agent is responsible, this is probably because it's *relevant* in some way to the agent's responsibility, and insofar as it does not seem relevant, it's unlikely to produce the desired judgement.

Note also that no *other* feature of the example seems to be doing any work in supporting our judgement that the agent is responsible. Once we make it clear that her alternatives are irrelevant, this destroys that judgement, even keeping the other features of the case constant.

Pereboom's example differs from McKenna's. It uses a moral 'buffer' that the agent would need to cross in order to avail himself of any relevant alternative.⁷

2.2 Pereboom's FSC

Pereboom aims to construct an example with the following three features:

1. The cue for intervention is a *necessary*, not sufficient condition, not for the action the agent actually performs, but for the agent's availing himself of any robust alternative possibility (without the intervener's device in place).

⁷For a more detailed discussion of this example, please see Elzein (2013).

2. The cue for intervention is not itself a robust alternative possibility.
3. The absence of the cue for intervention in no sense causally determines the action the agent actually performs.

These features are allegedly present in the following example:

Tax Evasion (2):

Joe is considering whether to claim a tax deduction for the substantial local registration fee that he paid when he bought a house. He knows that claiming the deduction is illegal, that he probably will not be caught, and that if he is, he can convincingly plead ignorance. Suppose he has a very powerful but not always overriding desire to advance his self-interest regardless of the cost to others, and no matter whether advancing his self-interest involves illegal activity. Crucially, his psychology is such that the only way that in this situation he could *fail* to choose to evade taxes is for moral reasons. His psychology is not, for example, such that he could fail to choose to evade taxes for no reason or simply on a whim. In addition, it is causally necessary for his failing to choose to evade taxes in this situation that he attain a certain level of attentiveness to these moral reasons. He 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, Joe could, with his libertarian free will, either choose to evade taxes or refrain from so choosing (without the interveners device in place). More generally, Joe is a libertarian free agent. But to ensure that he choose to evade taxes, a neuroscientist now implants a device, which, were it to sense the requisite level of attentiveness, would electronically stimulate his brain so that he would choose to evade taxes. In actual fact, he does not attain this level of attentiveness, and he chooses to evade taxes while the device remains idle (Pereboom 2003, p. 193; 2007a, pp. 90-91; 2009, p.113. Cf. 2001; 2000).

In this scenario, Pereboom maintains that Joe is morally responsible for choosing to evade his taxes despite the fact that there were no morally significant alternatives present.

Step 1:

Again, we must start by asking what the main differences are between this example and our simple case. In Simple Case 1, Brenda was faced with a choice between sipping her coffee and not sipping her coffee. In contrast, the choice that Joe faces is between bothering to attend to what is morally at stake before he makes a potentially morally dubious decision, and not paying attention to what is moral at stake prior to making his decision.

The most obvious difference again relates to the *sort* of alternative included. Unlike “taking a sip from what you believe to be an ordinary cup of coffee”, “attending to the moral reasons” looks as if it has obvious moral importance, so we might think the cue for intervention *does* look like a robust alternative. Pereboom argues that this could not possibly count as robust in this scenario, so let’s move on to step two, and examine this argument.

Step 2:

Pereboom argues that an alternative will not count as “robust” unless it is *likely* to exempt an agent from blame. This requires that the odds of exemption if the alternative is chosen must be high *enough*. The exact degree of probability required for an alternative to qualify as “likely” to be exempting is impossible to determine, but there will nonetheless be cases in which the probability is suitably low, and in those cases it will be clear that this alternative is not likely to exempt the agent from blame. Such alternatives will be insignificant (Pereboom 2009, pp. 111–112).

Does this plausibly establish that Joe’s alternative is non-robust? In this example, if Joe does not properly attend to the moral reasons, he will make avoiding blame *literally impossible*. We may wonder whether Joe *understands* that failing to attend to the reasons will reduce his capacity to make morally competent choices in this way, and hence whether he ought to recognise that he has a moral *duty* to attend to these reasons. If so, then it will not be at all obvious that the alternative is non-robust. The problem seems to be that the example leaves the nature of Joe’s neglect *ambiguous*: Whilst we recognise that we should all strive to be sensitive to the moral demands of our situation, we do not know whether Joe’s failure on this score constitutes a deliberate neglect of his duty, or whether it’s an irrelevant accident. Finally, we can easily remove this ambiguity by sharpening the example.

Step 3:

Firstly, let’s make it much more obvious that Joe fails to attend to the moral reasons deliberately, knowing that this will render it impossible for him to avoid later wrongdoing. Joe knows that he is about to deliberate about how to fill in his tax return. Joe recalls that he is often bogged down by nagging moral concerns when he makes such decisions, and he would rather these did not influence his choice. He therefore embarks on a deliberate conscious effort to completely ignore any moral considerations when making his choice. He distracts himself from these via any means available specifically in order to ensure that they do not influence his choice. Suppose he also knows full well that by adopting this policy, he will render it literally impossible for him to avoid cheating his taxes. Of course, at any time, he could drop this policy (thereby crossing the buffer), but he purposefully chooses not to precisely because he does not want such moral considerations to influence him.

It’s important that Joe is able to do this without thereby reflecting on the moral considerations he seeks to avoid, and hence triggering the buffer. But distracting oneself from thinking about moral considerations need not involve reflecting on the *content* of any concrete moral concerns one has. Perhaps he knows that if he starts contemplating the underfunding of public services, recent government cuts to disability benefits, low wages in state sector jobs, etc., this will prompt him to start contemplating the morality of tax evasion, so he purposefully seeks to keep his mind off these topics—perhaps he switches off the TV and hides the newspaper in order to ensure that the required level of attentiveness to such things does not occur in the first place. The important point, for our purposes, is that his effort must be deliberate, and not merely unknowing or accidental, and he must be broadly aware that by choosing

this course of action, he is removing all possibility of doing what he understands he morally ought to do.

While Joe certainly looks morally responsible in this example, it does not seem at all obvious that his alternatives are irrelevant. Rather, he seems to be responsible precisely because he took a deliberate course of action that he fully understood would make it impossible for him to refrain from making a morally bad choice.

So what if we instead sharpen in the opposite direction? Suppose, like Brenda, he is absolutely predetermined to commit the wrongdoing (in this case, illegally claiming a tax deduction), apart from his ability to make one genuinely free choice in the middle of his deliberation: He can freely decide whether or not to sip his coffee. If he takes a sip from his coffee, which has been spiked with heavy sedatives unbeknownst to him, he will be too sleepy to fill in the form until after the deadline has passed. He has no idea that his coffee is spiked, and hence, this truly is a morally irrelevant alternative. In contrast to “attending to the moral reasons”, it’s unambiguously clear that choosing whether to take a sip from his coffee has no moral importance, as far as he can reasonably suspect. And let’s suppose he does not choose to drink any coffee, hence rendering it literally impossible for him to avoid claiming the tax deduction. In this case, it does not seem at all obvious that he is morally responsible. Rather, like Brenda, it seems as if he did not have adequate opportunity to avoid wrongdoing. If he did not seem responsible in a purely deterministic context, a context with only this alternative present will fare no better.

Of course, we can also stipulate that the buffer might be crossed at any time. Perhaps right up until he sends off the relevant forms, he could take a sip from his coffee. Moreover, we can also suppose that crossing this buffer does not guarantee that he avoids wrongdoing. Perhaps if he *does* unwittingly ingest the sedatives, he may well go on to eat a Ginger Nut biscuit, which, unbeknownst to him, is spiked with stimulants, which would counteract the sedatives, enabling him to fill in the forms after all. But since he never crosses the buffer, and he appears not to be responsible for his failure to cross it, it does not seem at all clear that this additional detail (which is equally morally irrelevant as far as he knows) adds anything much to his responsibility.

Once again, to the extent that we are given reason to accept that the alternative really *is* irrelevant, it no longer looks at all obvious that it is fair to hold an agent responsible who has *only* that alternative available.

Again, the important differentiating feature, as far as our judgement that Joe is responsible goes, has to do with the sorts of alternative in play. When we make it clear that these *are* irrelevant, this undermines that judgement, clearly showing that none of the other features of the scenario are doing any work in supporting it. Again, none of the other features even look like plausible candidates for this role. And once again, we should ask the following question: Why should this feature make such a difference to the agent’s responsibility when it’s also supposed to be irrelevant to it? If including this feature is able to supply the otherwise lacking intuition that the agent is responsible, it is probably *relevant* to his responsibility.

That is all I want to say about FSCs that include a choice between morally undifferentiated alternatives. In the following section, I will examine FSCs which involve alternatives that are alleged to be morally irrelevant because they are inaccessible.

3 FSCs and Inaccessible Alternatives

Not all alternatives are accessible to agents. Being struck by lightning is an example of an alternative whose presence would have no bearing on an agent's blame. Some FSCs are intended to show that an agent might nonetheless appear responsible when only alternatives of this nature are present. As I said earlier, there are various examples that take this form, and I think that something similar can be said in each case, but I do not have space to consider all of them. Instead, I will focus on the sort of examples offered by Fischer and Stump, though I will later argue that my conclusion can be generalised to other examples of this sort.

3.1 Fischer and Stump's FSC

Fischer provides the following FSC:

Jones has left his political decision until the last moment, just as some diners leave their decision about what to order at a restaurant to the moment when the waiter turns to them. In any case, Jones goes into the voting booth, deliberates in the "normal" way and chooses to vote for the Democrat. On the basis of this choice, Jones votes for the Democrat. Unbeknownst to Jones, he has a chip in his brain that allows a very nice and highly progressive neurosurgeon (Black) to monitor his brain. The neurosurgeon wants Jones to vote for the Democrat, and if she sees that Jones is about to do so, she does not intervene in any way—she merely monitors the brain. If, on the other hand, the neurosurgeon sees that Jones is about to choose to vote for the Republican, she swings into action with her nifty electronic probe and stimulates Jones's brain in such a way to ensure that he chooses to vote for the Democrat (and goes ahead and votes for the Democrat). Given the set-up, it seems that Jones freely chooses to vote for the Democrat and freely votes for the Democrat, although he could not have chosen or done otherwise (Fischer 2007, p. 58. Cf. 1999; 2003; 2006).

Stump's version is similar. Here, the neurosurgeon knows that neuron firings a, b, c, always result in one choice whereas neuron firings x, y, z, always result in another. Specifically, these are early neural signals that would culminate in a decision unless Black intervenes (Stump 1999a, pp. 303-305; 2003, p. 140). So let's proceed once more through the same three-step procedure:

Step 1:

There is a strikingly obvious difference between this sort of example and Simple Case 2. In that case, the agent's alternative is clearly not accessible to her. She has no control at all over whether or not she is struck down by lightning. If it does happen to her, it would happen to her *against her will*. In the Fischer-Stump case, the agent's alternative is not so obviously inaccessible: It is possible that he exhibits an alternative pattern of neuron-firings, which would culminate in an alternative decision. Such an alternative, if it occurs, certainly will not normally be described as something that happens to him against his will, whether he likes it or not. It seems far less obvious

then, that it is a clear example of something that merely happens to the agent, and over which he exercises no control.

This example involves the neuron-firings that culminate in a decision, and it's very natural to assume that these might well be accessible in a way that being struck by lightning is not. The obvious thing a casual observer might be inclined to think is that these early neuron-firings *are* part of deciding or willing, and hence are something the agent *does* as opposed to something that merely *happens* to him. Once again, however, we are offered an argument for the conclusion that these events are in fact irrelevant. So let's move on to the second step of our procedure and examine this argument.

Step 2:

Both Stump and Fischer acknowledge that there would be a problem here if the neural patterns in question constituted "trying to decide" to vote Democrat, or "beginning to decide" or something like that. This would make it unclear whether the exhibiting of the alternative pattern is genuinely inaccessible. Stump argues that there is a "one-many" correlation between neural events and choices, and that until all of the neuron-firings are complete, nothing like a decision has occurred. In this case, what we really have is an involuntary brain event (Stump 1999a, p. 304; b, p. 417). This argument is strongly advocated by Fischer too, who also maintains that these early signals may be safely regarded as involuntary brain events (Fischer 1999, pp. 115-117; 2006, pp. 194-195).

The problem, however, is that many of us are strongly inclined to suppose (rightly or wrongly) that we *can* initiate our own decisions, and hence, are likely to suppose that this alternative is accessible to Jones. And of course, The Fischer-Stump argument requires us to accept that our own decisions are the sorts of thing that merely *happen* to us; if we cannot have any voluntary control until *after* the choice is completed, then clearly our decisions could not be something we are able to *bring about*. This is a highly controversial claim. Whether or not the argument is convincing, it seems the example trades on the fact that our judgements about this sort of thing are conflicting at best. Finally, since the precise relevance of the alternatives is unclear in this example, let's once again see how sharpening it affects our judgements.

Step 3:

Firstly, suppose that Jones's choice is entirely determined, except for one genuinely chancy event that might occur: It might be that during his deliberation, he is involuntarily subject to some sort of random epileptic seizure that disrupts his neural activity. Suppose that Jones is otherwise causally determined to vote Democrat, save for the possibility that this chancy epileptic seizure occurs. If he is struck by this seizure, the disruption this causes to his brain activity would make him vote Republican, unless Black's device should intervene forcing him to vote Democrat. Now suppose that that Black's device monitors his brain and remains idle unless it should detect that such a seizure has occurred. As it happens, Jones does not suffer the chancy seizure and, hence, cannot possibly do anything other than vote Democrat.

In this example, it's far from obvious that Jones can be held responsible for his choice. After all, his choice is entirely determined (much like Brenda's choice in our

Simple Case 2), save for the chance of something happening to him that is entirely outside of his control. Once we make sufficiently clear that this is the only possible alternative present, and that it really is something that, if it occurs, would merely *happen* to him, as opposed to being something that he *does*, it no longer seems at all obvious that he is responsible. At least, if we are inclined to doubt his responsibility in a purely deterministic context, a context that includes only *this* sort of alternative seems to fare no better.

So what about if we sharpen it in the opposite direction? Suppose that what we have are not involuntary brain events at all, but rather that these brain events constitute efforts of will (or perhaps the initiation of such efforts of will), and Jones has the power to deliberately bring about either one, where one would result in his voting Democrat, and the other would result in his voting Republican (and the decision about whether to manipulate Jones will be based on which of these he does). Again, this seems to rescue the intuition that Jones is responsible, but it does so only because we have reintroduced accessible alternatives.

Once again, the important difference is in the nature of the alternatives, as can be seen when we sharpen the example. Once again, there are no other features of the example that even look like plausible contenders for the source of this judgement. As with the previous cases, we should ask the same question: If changing the nature of the alternatives in this way suddenly transforms the agent into a responsible one (in contrast to our simple case), is it not likely that these alternatives are relevant to his responsibility?

3.2 Can We Generalise This Conclusion?

Since I have only singled out one example of an FSC in which the alternatives are said to be inaccessible to the agent, it will be worth saying something about some of the others. The Stump-Fischer example is a good one to begin with, since they are particularly clear about where the alternatives are and why they are thought to be inaccessible. But similar arguments appear to be at work in other versions of the example. I think we have good grounds to suppose that my conclusion can be generalised.

There is one feature shared by *all* of these examples, which differentiates them from the simple case: The alternatives, which are supposed to be inaccessible, are always events located within the agents' brains, and closely related to their decisions and deliberation. In contrast, for the simple case, I used lighting strikes, which are much more obviously inaccessible.

There is always some difference in these examples between the actual scenario and the alternative one, although the nature of this difference changes from one version of the example to another. But there must always be *some* difference between the actual and alternative scenarios, since we must maintain that the agent actually acts of his own accord, where the same agent would have been manipulated instead were the alternative scenario to occur. This is why it has proved so difficult to eliminate *all* alternative possibilities from FSCs.

In Hunt's FSC, it is the difference between an agent's decision occurring only due to a neuron blockage and that same decision occurring irrespective of the neuron

blockage (Hunt 2000). In Robb and Mele's FSC, it is the difference between a choice that results from the agent's indeterministic brain processes, and a choice that occurs instead because of an extraneous deterministic process, that can only be stopped by the indeterministic process (Mele and Robb 1998, 2003). In Widerker's version, it is the difference between an agent choosing wrongdoing irrespective of a brain malfunction which ensures that it is impossible for him to arrive at a blameless decision, and the agent choosing the wrongdoing only *because* this brain malfunction means it is impossible for him to arrive at a blameless decision (Widerker 2006, 2009). And of course, there may be different versions with alternatives that are differentiated in other ways.

In *all* of these FSCs, it's clearly the fact that they include alternatives that involve events which are located in the brain and involved in the agent's choices and deliberation that makes them fare better than the simple case when it comes to evoking the judgement that the agent is responsible (in contrast to the simple case, where the alternative involves being passively subject to a lightning strike). Generally, it is some sort of neural event that would trigger the counterfactual mechanism if it is not going to culminate in the desired decision, where that neural event is closely involved in the agent's choice.

It is very clear why this sort of feature is popular for use in these examples. This is one area where our judgements tend to be confused. When we think about making a decision, we think of this as something that we *do* rather than something that merely *happens* to us. When we think about the brain events that culminate in a decision, however, it's possible to argue that we are merely passively *subject* to these. Of course, if the brain events that culminate in decisions simply *happen to us*, then *initiating* a choice cannot be something we actively *do*, but must also be something that simply *happens* to us. We can easily be persuaded to think of brain events in this way, but this is not the way we typically think of choices. However, we also know (at least in the abstract) that these are simply one and the same thing.

We cannot, however, have it both ways. The examples work by playing on our everyday intuitions (which tell us that decisions are something we do and not something that happens to us) and then providing arguments which, if we *accept* them, would destroy the very intuitions that the example relies on (by implying that these same decisions are made up of neural events, all of which simply happen to us whether we like it or not). To the extent that we are convinced by the arguments, we are also inclined to doubt that the agent is responsible. To the extent that we are not convinced, we can sustain our intuition that the agent is responsible, but at the cost of including alternatives that still look highly robust. But it's very difficult to sustain both of these judgements *at once*, which is what the neo-Frankfurtian requires.

That is not to say, of course, that there could be no good independent arguments to the conclusion that we ought to accept both that the neural events in question are inaccessible, and hence irrelevant, and that we ought to regard the agent as responsible anyway, irrespective of any intuitive judgements to the contrary. If, however, we can construct such an argument, this would of course render the FSCs themselves redundant: Their force would depend entirely on that of this independent argument, which, if it succeeds, would show the irrelevance of alternative possibilities on its own.

4 Possible Responses

Fischer has claimed that weakly successful FSCs provide a *plausible* argument against the leeway view, and may contribute to a broader holistic argument against it (Fischer 2013). But as he says himself: "...if it could be argued that in every proposed FSC there is at least one robust alternative possibility, then this would be quite a different matter. But this is where the argument would need to focus, and thus far I have not been convinced by any such argument" (Fischer 2013, p. 57). This is precisely the challenge I have tried to take up here.

I claim that (a) none of the examples include only alternatives that are clearly and unambiguously morally insignificant, (b) when alternatives *are* made clearly insignificant, it is not at all intuitively obvious that the agent ought to be regarded as responsible (at the very least, if we think it would be question-begging to make such an assertion in purely deterministic contexts, these scenarios fare no better), and (c) that it is probably *impossible* to construct an example in which it is obvious that the agent is responsible despite the fact that any alternatives present are clearly insignificant. There may of course still be a good case against the leeway view, but I hope to have shown that FSCs make no significant contribution to it.

But even if present FSCs fail, the neo-Frankfurtian may try to counter my claim that it is not even *possible* to construct an effective FSC that includes only clearly irrelevant alternatives. It is worth exploring some lines of argument on the basis of which such optimism might be supported.

I have throughout aimed to compare the sorts of alternative present in "simple cases" (where they are obviously non-robust) with those in FSCs (where I have argued that it's generally at best ambiguous whether the alternatives are robust). If the agent in our FSC appears responsible when the agent in our simple case does *not*, then there must be some feature that differentiates these cases and that accounts for this disparity in judgements. Let's call this feature (whatever it is), Feature X.

In all of the popular present FSCs, the Feature X relates directly to the nature of the alternatives included. I have suggested that this is problematic. Here, I have relied on the following sort of inference:

1. Changing the nature of the alternatives in play (and nothing else) makes all the difference with regard to the agent's responsibility.
2. The alternatives in play are probably relevant in some way to the agent's responsibility.

This implies that if Feature X is related to the nature of the alternatives, then we cannot maintain that these are irrelevant to the agent's responsibility, and hence cannot maintain that our example weakly succeeds. A neo-Frankfurtian may, however, reject this inference. She might provide a further iteration of the same arguments that are used to undermine the claim that alternatives *per se* are relevant to moral responsibility.

The original argument involved conceding that there are alternative possibilities present in all of the examples in which it is clear that the agent is responsible, but insisting that these alternatives are doing no work in actually *supporting* this moral

responsibility (a point that would be supported by weakly successful FSCs). Perhaps the neo-Frankfurtian will be willing to further weaken her claim and then run this same sort of argument again: That is, she may concede not only that there must be alternative possibilities, but also that there must be alternative possibilities *of a very specific nature* present in all of the examples in which it is clear that the agent is responsible. But she may insist that alternatives *of this very specific nature* are doing no work in actually *supporting* this moral responsibility.

Is this move feasible? I doubt it. The neo-Frankfurtian view (that alternatives *per se* are irrelevant) is much more plausible if we can also maintain that the *specific nature of the alternatives* does not matter. The idea that the alternatives are irrelevant *combined with* the claim that it is not alternatives *per se*, but alternatives of a *very specific nature* that must be present is much less compelling.

At the very least, it would help if we could actually identify the relevant feature that the alternatives would need to have in a possible successful FSC and could show that this feature (a) is morally insignificant and (b) might for some reason plausibly be twinned to the agent's responsibility despite its moral insignificance. But in this respect, this new (neo-neo-Frankfurtian?) view seems to be on a much weaker footing than the original neo-Frankfurtian view. There are fairly intuitive reasons why we might suspect that mere indeterminism could be twinned with responsibility (or at least with common judgements about it) despite having no supporting role. For example, it could be a very broad and basic precondition of sourcehood, or it could guarantee simple independence from certain sorts of prior influence, or it might just be that the long historical assumption that there is an opposition between responsibility and determinism colours our judgements. In contrast, it's hard to see how a convincing case could be made for the idea that the presence of alternatives *of very specific sorts* might be twinned with responsibility despite the fact that alternatives in general have no supporting role.

Further doubt is cast on this idea by the very fact that *none* of the popular present FSCs include alternatives that have this mysterious nature, and it's hard to even imagine what it might be. In *every* present FSC, the alternatives differ from those of the simple cases just in virtue of being much more morally robust. Furthermore, we can play around with sharpening the examples, and we find that our judgements shift *precisely* in line with how obvious the robustness of the alternative is.

Of course, I cannot prove that it's impossible to construct a convincing argument of this sort and hence to construct new FSCs that *do* succeed (albeit even more weakly!) But I suspect it will be a significant challenge.

There may, however, be a second strategy that the neo-Frankfurtian could adopt in order to establish that convincing FSCs are, at least in theory, *possible* to construct. In relation to all of the present FSCs, I have suggested that Feature X relates directly to the nature of the alternatives included. As we have just seen, it is this that causes the problem.

But at this point, we may wonder whether Feature X *must* relate to the nature of the alternatives. In present FSCs, there are no *other* features, unrelated to the nature of the alternatives, that even look like plausible *candidates* for being the source of the judgement that the agent is responsible. But perhaps, it may be possible (at least in theory)

to find some such feature. Our hypothetical example would then be well equipped to show that an agent may be responsible despite the fact that any alternatives present are *clearly* insignificant.

There are, however, three questions that I think we ought to ask here:

Firstly, how exactly does Feature X produce the judgement that the agent in our example (unlike the agent of our simple case) is morally responsible? Perhaps Feature X is just *obviously* sufficient for moral responsibility. Or perhaps there is some good *argument* to the conclusion that Feature X ought to be regarded as sufficient for moral responsibility. Either way, if we can establish this much, it is not at all obvious why we would need to prop our argument up with FSCs. Since, *ex hypothesi*, Feature X is unrelated to the presence or nature of any alternatives present in the example, there is no reason to suppose this feature could not be present both in simple cases and in purely deterministic setups. If Feature X can be shown to be (or is obviously) sufficient for moral responsibility, then we do not need to appeal to FSCs to show the irrelevance of alternative possibilities.

This brings me to my second question: Why bring scenarios involving counterfactual intervention into the dispute at all? If Feature X is sufficient for responsibility, and it does not relate in any way to alternative possibilities, it is unclear why it should help us to think about it in relation to these alternatives, which everyone must agree from the outset are completely irrelevant. Given that we must all agree that the alternatives are irrelevant, it should hardly be surprising that examples which centre around such alternatives do not tell us very much.

Finally, why should we suppose this feature exists? Essentially, FSCs are supposed to establish precisely that there could be a Feature X that does not require alternative possibilities. Without a successful FSC, we have no reason to even *suspect* that this is true (unless of course, we could offer an independent argument to this conclusion, which would render FSCs themselves redundant anyhow).

I am optimistic about the prospects of arguing plausibly for sufficient conditions of moral responsibility that do not require alternative possibilities. In *Free Will and the Concept of a Person*, Frankfurt (1971) himself does exactly that. But such arguments do not derive their force from FSCs: Rather, the FSCs depend for their force on the success of such independent arguments, and if the independent argument succeeds, the FSCs are redundant.

5 Conclusion

It should now be possible to see that the problem for the neo-Frankfurtian runs deeper than just the challenge of constructing an example that only includes morally irrelevant alternatives. If the example is also to have the required intuitive impact, we will have to overcome some serious difficulties.

Suppose that Feature X relates to the nature of the alternatives in the example. I hope to have shown that this provides us with very good reason to suspect that these alternatives are relevant in some way to the agent's responsibility. Moreover, this is precisely how *all* of the present popular FSCs work: They include alternatives that are

at least of borderline relevance, and when we sharpen the examples, we can indeed confirm that the agents only appear to be responsible insofar as they appear to have *robust* alternatives available to them.

If, in contrast, Feature X is unrelated to any alternatives in play in the example, I hope to show that the FSC setup quickly becomes irrelevant to the dispute and that FSCs themselves become redundant.

But even for those who are not persuaded that it is genuinely *impossible* to construct a weakly successful FSC, I hope to at least challenge the whole method (common to all of the popular *present* FSCs) of focusing on increasingly complex examples with increasingly obscure alternatives, as opposed to focusing on simple examples involving *obviously* irrelevant alternatives. If it does not seem as if the agent is responsible when the alternatives are *obviously* irrelevant, why should we suppose that we would get any closer to the truth by instead focusing on alternatives that are of obscure or borderline relevance? Why should muddying the waters in this way help us to see what really matters for moral responsibility?

If we want to find out whether agents appear to be responsible in scenarios where there are only irrelevant alternatives present, we should clearly expect examples in which the alternatives are *obviously* irrelevant to be most informative here. Shifting our attention instead to examples in which there is a complex *dispute* to be had about the relevance of the alternatives seems at best a bizarre move, and at worse deliberately misleading.

In fact, when we consider the difference that is made by really spelling out the relevance or irrelevance of alternatives, we find that the more robust the alternative possibilities are, the more obvious it seems that the agent is morally responsible, and the more morally insignificant the alternative possibilities are, the more doubtful it seems that the agent is responsible. If anything then, considering examples of this kind provides a great deal of intuitive *support* to the leeway view. It very much appears to be the presence of robust alternatives (as opposed to mere indeterminism) that underpins our intuitive judgements about moral responsibility.

Of course, it may well be that there are *other* problems with the leeway view. Perhaps it will be shown that there are good arguments against all purported API principles, or that robust alternative possibilities always come with responsibility-undermining randomness, or that the very idea of being able to initiate alternative choices commits us to incoherent nonsense. Similarly, perhaps there will be good positive arguments in favour of a compatibilist view or source incompatibilist view, which would undermine any case in favour of the leeway view. What I have said leaves all of these possibilities open. I have also left open that it may be possible to argue that deterministic FSCs are not question-begging. I believe, however, that the most convincing case has always come from indeterministic FSCs. And I hope to have shown that these do not help to further the dispute.

And finally, I do not intend to say that FSCs do not help to illuminate the free will dispute at all. Many compatibilists traditionally also accept some version of PAP, and it *certainly* does not beg the question against the *compatibilist* to presuppose a deterministic setting, regardless of whether it begs the question against the incompatibilist.

In this respect, FSCs may well help a compatibilist (or in theory, a source incompatibilist) to bring *her own* commitments into sharper focus, even if they do not help to refute API principles.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Alvarez, M. (2009). Actions, thought-experiments and the ‘principle of alternate possibilities’. *Australasian Journal of Philosophy*, *1*, 61–81.
- Ekstrom, L. (2003). Free will, chance, and mystery. *Philosophical Studies*, *2*(1), 153–180.
- Elzein, N. (2013). Pereboom’s Frankfurt case and derivative culpability. *Philosophical Studies*, *166*(3), 553–573.
- Fischer, J.M. (1999). Recent work on moral responsibility. *Ethics*, *110*(1), 93–139.
- Fischer, J.M. (2003). Frankfurt-style compatibilism. In Watson, G. (Ed.) *Free will*. 2nd edn. (pp. 190–211): Oxford University Press.
- Fischer, J.M. (2006). Frankfurt-type examples and semi-compatibilism. In Kane, R. (Ed.) *The Oxford handbook of free will* (pp. 281–308): Oxford University Press.
- Fischer, J.M. (2007). Compatibilism. In Kane, R., Fischer, J.M., Pereboom, D., & Vargas, M. (Eds.) *Four views on free will* (pp. 44–84): Blackwell Publishing.
- Fischer, J.M. (2010). Frankfurt cases: the moral of the stories. *Philosophical Review*, *119*, 315–336.
- Fischer, J.M. (2013). The Frankfurt-style cases: philosophical lightning rods. In Caouette, J., & Haji, I. (Eds.) *Free will and moral responsibility*, chapter 2, (pp. 43–57): Cambridge Scholars Publishing.
- Frankfurt, H.G. (1969). Alternate possibilities and moral responsibility. *The Journal of Philosophy*, *66*(23), 829–839.
- Frankfurt, H.G. (1971). Freedom of the will and the concept of a person. *The Journal of Philosophy*, *68*(1), 5–20.
- Franklin, C. (2011). Neo-Frankfurtians and buffer cases: the new challenge to the principle of alternative possibilities. *Philosophical Studies*, *152*, 189–207.
- Ginet, C. (1996). In defence of the principle of alternate possibilities: why I don’t find Frankfurt’s argument convincing. *Philosophical Perspectives*, *10*, 403–417.
- Goetz, S. (2005). Frankfurt-style counterexamples and begging the question. *Midwest Studies in Philosophy*, 83–105.
- Goetz, S. (2008). *Freedom, teleology, and evil*. Continuum.
- Haji, I., & McKenna, M. (2004). Dialectical delicacies in the debate about freedom and alternative possibilities. *The Journal of Philosophy*, *101*, 299–314.
- Haji, I., & McKenna, M. (2006). Defending Frankfurt’s argument in deterministic contexts: A reply to Palmer. *The Journal of Philosophy*, *103*, 363–372.
- Hunt, D.P. (2000). Moral responsibility and unavoidable action. *Philosophical Studies*, *97*, 195–227.
- Hunt, D.P. (2005). Moral responsibility and buffered alternatives. *Midwest Studies In Philosophy*, *29*(1), 126–145.
- Kane, R. (1985). *Free will and value*. State University of New York Press.
- Kane, R. (1996). *The significance of free will*. Oxford University Press.
- Kane, R. (2000). The dual regress of free will and the role of alternative possibilities. *Philosophical Perspectives*, *14*, 57–79.
- Kane, R. (2002). Some neglected pathways in the free will labyrinth. In Kane, R. (Ed.) *The Oxford handbook of free will, chapter 18* (pp. 406–37): Oxford University Press.
- Kane, R. (2004). Agency, responsibility, and indeterminism: reflections on libertarian theories of free will. In Campbell, J.K., O’Rourke, M., & Shier, D. (Eds.) *Freedom and determinism*, chapter 3, (pp. 77–78): MIT Press.

- Kane, R. (2007a). Libertarianism. In Kane, R., Fischer, J.M., Pereboom, D., & Vargas, M. (Eds.) *Four views on free will*, chapter 1, (pp. 5–43): Blackwell Publishing.
- Kane, R. (2007b). Response to Fischer, Pereboom and Vargas. In Kane, R., Fischer, J.M., Pereboom, D., & Vargas, M. (Eds.) *Four views on free will*, chapter 5, (pp. 166–183): Blackwell Publishing.
- Leon, F., & Tognazzini, N.A. (2010). Why Frankfurt-examples don't need to succeed to succeed. *Philosophy and Phenomenological Research*, 80(3), 551–565.
- McKenna, M.S. (2003). Robustness, control, and the demand for morally significant alternatives: Frankfurt examples with oodles and oodles of alternatives. In McKenna, M., & Widerker, D. (Eds.) *Moral responsibility and alternative possibilities: essays on the importance of alternative possibilities*, chapter 11, (pp. 201–218): Ashgate Publishing Ltd.
- Mele, A.R., & Robb, D. (1998). Rescuing Frankfurt-style cases. *The Philosophical Review*, 107(1), 97–112.
- Mele, A.R., & Robb, D. (2003). Bbs, magnets and seesaws: the metaphysics of Frankfurt-style cases. In Widerker, D., & McKenna, M. (Eds.) *Moral responsibility and alternative possibilities: essays on the importance of alternative possibilities*, chapter 7, (pp. 107–126): Ashgate Publishing Ltd.
- Moya, C. (2006). *Moral responsibility: the ways of scepticism*. Routledge.
- Moya, C. (2007). Moral responsibility without alternative possibilities? *The Journal of Philosophy*, 104, 475–486.
- Moya, C. (2011). On the very idea of a robust alternative. *Critica*, 43(128), 3–26.
- Otsuka, M. (1998). Incompatibilism and the avoidability of blame. *Ethics*, 108, 685–701.
- Pereboom, D. (2000). Alternative possibilities and causal histories. *Philosophical Perspectives*, 14, 119–137.
- Pereboom, D. (2001). *Living without free will*. Cambridge University Press.
- Pereboom, D. (2003). Source incompatibilism and alternative possibilities. In Widerker, D., & McKenna, M. (Eds.) *Moral responsibility and alternative possibilities: Essays on the importance of alternative possibilities*, chapter 10, (pp. 185–199): Ashgate Publishing Ltd.
- Pereboom, D. (2007a). Hard incompatibilism. In Kane, R., Fischer, J.M., Pereboom, D., & Vargas, M. (Eds.) *Four views on free will*, chapter 3, (pp. 85–125): Blackwell Publishing.
- Pereboom, D. (2007b). Response to Kane, Fischer and Vargas. In Kane, R., Fischer, J.M., Pereboom, D., & Vargas, M. (Eds.) *Four views on free will*, chapter 7, (pp. 191–203): Blackwell Publishing.
- Pereboom, D. (2009). Further thoughts about a Frankfurt-style argument. *Philosophical Explorations*, 12(2), 109–118.
- Shabo, S. (2010). Uncompromising source incompatibilism. *Philosophy and Phenomenological Research*, 80(2), 349–383.
- Steward, H. (2009). Fairness, agency and the flicker of freedom. *Noûs*, 43(1), 64–93.
- Stump, E. (1999a). Alternative possibilities and moral responsibility: the flicker of freedom. *The Journal of Ethics*, 299–324.
- Stump, E. (2003). Moral responsibility without alternative possibilities. In Widerker, D., & McKenna, M. (Eds.) *Moral responsibility and alternative possibilities: essays on the importance of alternative possibilities*, chapter 8, (pp. 139–158): Ashgate Publishing Ltd.
- Stump, E. (1999b). Dust, determinism, and Frankfurt: a reply to Goetz. *Faith and Philosophy*, 16(3), 413–422.
- Timpe, K. (2007). Source incompatibilism and its alternatives. *American Philosophical Quarterly*, 44(2), 143–155.
- Timpe, K. (2008). *Free will: sourcehood and its alternatives*. Continuum.
- van Inwagen, P. (1978). Ability and responsibility. *The Philosophical Review*, 87, 201–224.
- Widerker, D. (1995). Libertarianism and Frankfurt's attack on the principle of alternate possibilities. *Philosophical Review*, 104, 247–261.
- Widerker, D. (2006). Libertarianism and the philosophical significance of Frankfurt scenarios. *The Journal of Philosophy*, 103(4), 163–187.
- Widerker, D. (2009). A defence of Frankfurt—friendly libertarianism. *Philosophical Explorations*, 12(2), 87–108.
- Zagzebski, L. (2000). Does libertarian freedom require alternate possibilities? *Noûs*, 34, 231–248.
- Zagzebski, L. (2010). Foreknowledge and human freedom. In Taliaferro, C., Draper, P., & Quinn, P.L. (Eds.) *A companion to philosophy of religion*, chapter 56, (pp. 474–481): Wiley-Blackwell.