

On Fischer and Frankfurt-style Cases

William Simkulet

Abstract

Almost everyone believes that moral responsibility requires control; however, philosophers disagree about whether this control is compatible with universal causal determinism. Many philosophers argue that it is not, and to illustrate this intuition they turn to the principle of alternate possibilities (PAP) to demonstrate this incompatibility. According to PAP, for an agent to be morally responsible for her action, she must have been able to do otherwise. If our actions are causally necessitated by circumstances that occurred long before we were ever born, it wouldn't make sense to say we are responsible because we lack both alternate possibilities and control. Recently, compatibilists – starting with Harry Frankfurt – have attempted to construct counter-examples to PAP – Frankfurt-style cases – in which an agent is said to be morally responsible while lacking alternate possibilities. In “The Frankfurt Cases: The Moral of the Stories,” John Martin Fischer defends Frankfurt-style cases from what he calls “The Dilemma Defense.” Here I argue Fischer's defense fails.

On Fischer and Frankfurt-style Cases

Although most philosophers believe moral responsibility requires control, they disagree about what kind of control is sufficient for moral responsibility. Notably, there is disagreement between as to whether moral responsibility is compatible with universal causal determinism, where *universal causal determinism* is the theory that there is only one possible future and that future is completely causally necessitated by the laws of nature and circumstances of the distant past. *Compatibilists* believe moral responsibility is compatible with determinism, while *incompatibilists* argue that responsibility is fundamentally incompatible with determinism. To circumvent complex metaphysical and metaethical issues, many incompatibilists turn to the principle of alternate possibilities to illustrate this incompatibility:

Principle of Alternate Possibilities (PAP) – A necessary (but not sufficient) condition for agent *a*'s being morally responsible for *x* is that *a* could have done other than *x*.

Traditionally, both compatibilists and incompatibilists have found PAP to be a commonsense, intuitively plausible moral principle, yet incompatibilists argue that if PAP is true, moral responsibility is incompatible with universal causal determinism because if universal causal determinism is true, then there is only one possible future and no one can actually do other than they actually do – they lack genuine alternate possibilities.

In “Alternate Possibilities and Moral Responsibility,” Harry Frankfurt constructs a case intended to be an open-ended counter-example to PAP in which an agent – Jones – is said to (i) *uncontroversially* lack alternate possibilities and yet (ii) be *uncontroversially* morally responsible for his actions. To avoid begging the question about the compatibility of determinism and responsibility, Frankfurt intends the case to be *metaphysically neutral* such that regardless of one's metaphysical beliefs about causation, one would be inclined to believe Jones is uncontroversially responsible despite lacking alternate possibilities.

Critics argue that Frankfurt-style cases are not true counter-examples to PAP because they cannot be constructed in such a way as to show both (i) and (ii) – either Jones (~i) has alternate possibilities, or he (~ii) is not uncontroversially morally responsible; call this *the dilemma defense*.¹ In “The Frankfurt Cases: The Moral of the Stories,” John Martin Fischer sets out to defend Frankfurt-style cases from the dilemma defense, and attempts to show that (i) and (ii) can be true of such a case by building upon recent work in the field.² Here I will look at Frankfurt's original case, the dilemma defense, and Fischer's response, and argue that Fischer's response is unsatisfying because it abandons the open-endedness of Frankfurt-style cases. I will end by discussing what a satisfactory Frankfurt-style case would look like, and argue that the continuing inability of compatibilists to construct such a case suggests their theory is inconsistent with our metaethical intuitions about moral responsibility.

¹ See Robert Kane (1985, 1996); David Widerker (1995), Carl Ginet (1996); Stewart Goetz (2005).

² See Derk Pereboom (2008).

I. Frankfurt on PAP

In “Alternate Possibilities and Moral Responsibility,” Harry Frankfurt makes two substantive arguments against PAP:

- (1) PAP derives its intuitive force from the coercion principle, the coercion principle is false, and thus PAP lacks intuitive force.
- (2) It is possible to construct a case in which an agent (i) uncontroversially lacks alternate possibilities, and yet (ii) is uncontroversially blameworthy, so PAP is false.

The latter argument has been the focus of much recent work in philosophy, while the first has failed to garner much attention. Here I will look at both.

Frankfurt’s first argument contends that there is an intuitive connection between PAP and the coercion principle:

Coercion Principle (CP) – A necessary (but not sufficient) condition for agent *a*’s being morally responsible for *x* is that *a*’s action *x* is uncoerced.

One might appeal to a principle like CP to explain why it doesn’t make sense to hold coerced agents morally responsible in cases like this:

*Smith*¹: White kidnapped bank manager Smith¹’s family and told him that he will execute them in an hour unless he brings him a million dollars from the bank without alerting the police. Smith¹ believes White’s threat, loves his family, steals the money, and gives it to White.

Many people do not think it is appropriate to blame or punish Smith¹ for stealing the money under these circumstances. CP offers an explanation why – Smith¹ was coerced, and coercion undermines moral responsibility. But CP is false.

*Smith*²: White kidnapped bank manager Smith²’s cilantro plant and told him that he will set it on fire in an hour unless he brings him a million dollars from the bank without alerting the police. Smith¹ believes White’s threat, likes cilantro, steals the money, and gives it to White.

Even if Smith² was coerced by White (perhaps due to an atypical attachment to his cilantro plant), it doesn't make sense to say that he isn't responsible for what he does, and thus Smith² is a counter-example to CP. Smith² is uncontroversially blameworthy.

Furthermore, it makes sense to say that Smith¹ *is* responsible for his action – he took quite a risk to save his family's life and this is *prima facie* praiseworthy. James Rachels argues the right thing to do is the thing that one has the best reasons to do, and it is at least plausible to say that Smith¹ has more reasons to take the money than not to.³ It seems coerced agents can be either praiseworthy or blameworthy depending upon the reasons they act upon.

Frankfurt says:

Now the doctrine that coercion and moral responsibility are mutually exclusive may appear to be no more than a somewhat particularized version of [PAP]. It is natural enough to say of a person who has been coerced to something that he could not have done otherwise. And it may easily seem that being coerced deprives a person of freedom and of moral responsibility simply because it is a special case of being unable to do otherwise. The principle of alternate possibilities may in this way derive some credibility from its association with the very plausible proposition that moral responsibility is excluded by coercion (1969, 830–831).

It's not clear why Frankfurt believes the coercion principle is merely a particularized version of PAP as it seems coerced agents often can do otherwise. However, he does suggest that if one has been sufficiently coerced, then one lacks alternate possibilities because the alternatives are too horrific. This is quite suspect – surely people freely choose to act in horrific ways, so the mere fact that an alternate possibility is horrific does not preclude it from being a possibility (although it might make sense to say that choosing such a possible action would be immoral or irrational).

It is possible that an agent can be psychologically constituted such that certain scenarios trigger what Eddy Nahmias calls *bypassing*, the circumventing of an agent's normal moral deliberation processes in such a way that is inconsistent with moral responsibility. Nahmias's position is consistent with Frankfurt's – the existence of deterministic bypassing invites

³ See Rachels (2003).

confusion, potentially leading interpreters to the conclusion that determinism is responsibility-undermining, rather than bypassing. Note that for Nahmias, an agent is morally responsible for the choices that result from their normal moral deliberation, so if Smith¹ and Smith² were bypassed, neither would be responsible on his view.

To illustrate the falsity of CP, and undermine the intuitive force behind PAP, Frankfurt constructs a series of three cases:

1. Jones freely chooses to perform action A, and freely does so. As it so happens, a would-be coercer requests that he do A and threatens a harsh penalty for refusing. However, Jones ignored this threat and did A for his own reasons.
2. Jones freely chooses to perform action A. However, before he does A, a coercer requests that he do A and threatens a harsh penalty for refusing. Jones is stampeded, forgets his original free choice, and does as the coercer asks to avoid the penalty.
3. Jones freely chooses to perform action A. However, before he does A, someone requests that he do A and threatens a harsh penalty for refusing. Jones does not ignore the threat, and had he not already freely chosen to A, he would do A to avoid the harsh penalty. However, he freely As for his own reasons.

Frankfurt contends that at least one of these cases is a counter-example to CP, and by demonstrating the falsity of the coercion principle, Frankfurt believes he has undermined the appeal of PAP (834).

Anticipating criticism, Frankfurt constructs a final case designed to appease even the most ardent skeptics. In this case, he turns his attention away from CP towards PAP itself and attempts to construct a case in which an agent (i) uncontroversially lacks alternate possibilities, yet (ii) is uncontroversially morally responsible for his action. A concise version follows:

4. Black wants Jones to perform action A, and has a means of forcing Jones to do A. However, he would rather Jones freely choose to do A. Black has observed that in the past, whenever Jones chooses between an A and \sim A, he consistently twitches before the choosing \sim A. As it so happens, this time Jones doesn't twitch and freely chooses to A. Black doesn't intervene.

Frankfurt contends that in this case:

- (i) Uncontroversially, Jones could not have done otherwise. (Jones lacks alternate possibilities.)

(ii) Jones is uncontroversially morally responsible for his action.

If (i) and (ii) are true, this is a persuasive counter-example to PAP. For Frankfurt, it is not enough to argue that it makes sense to say Jones is responsible despite being unable to do otherwise; rather he means to show that this is *uncontroversially* the case. His goal is to construct a persuasive counter-example to PAP – one that will convince anyone, regardless of their metaethical intuitions regarding the compatibility of free will and determinism, that PAP is false.

Indeed, many find this case convincing. In “Incompatibilism and the Avoidability of Blame”, Michael Otsuka, convinced of Frankfurt's success, proposes an alternative to PAP, the *principle of avoidable blame* (PAB), according to which for one to be morally responsible, one must have had the possibility to act in a manner in which one would have been entirely morally blameless. Otsuka's principle strikes me as a plausible replacement for the coercion principle, but not PAP. PAB explains our intuitions in traditional coercion cases, but requires the same kind of alternate possibilities as PAP. Otsuka believes that Jones is wholly morally culpable for his action if he chooses it freely, but that he would have been entirely morally blameless had Black forced him to act. But this just is to say that Otsuka believes that Jones has alternate possibilities – one possibility in which he freely chooses to A, and one in which Black compels him to choose A. Jones has alternate possibilities, what he lacks are *alternate outcomes*; this is to say that Frankfurt's case suggests Jones can freely A or be forced to A by Black. In either possibility, the outcome is the same – Jones As; what differs is how Jones As (freely or by compulsion) and how responsible he is for it (responsible or not responsible).

II. Fischer on the Dilemma Defense

In “The Frankfurt Cases: The Moral of the Stories”, John Martin Fischer contends the primary threat to the legitimacy of Frankfurt cases comes from the *dilemma defense*. According to this argument, either:

(~i) Jones could have done otherwise.

or

(~ii) Jones is not uncontroversially morally responsible.

The argument goes as follows – either Jones's actions are (a) undetermined or (b) determined. If (a), then Black cannot know what he is going to do before he does it, and thus Black cannot cut off alternate possibilities. If (b), then Jones is not uncontroversially morally responsible because there is

substantial controversy over whether determinism is compatible with moral responsibility. Fischer calls (a) the *indeterministic horn* of the argument, and (b) the *deterministic horn*, and offers an argument against each. I briefly describe each horn, then discuss Fischer's reply.

A. The Indeterministic Horn

Early criticism of Frankfurt cases focused on critiquing the method by which Black is said to know what Jones will do.⁴ Frankfurt proposes that Black has discovered that Jones has a tell – a twitch that indicates how he will act in the future. Critics contend this account fails to show both (i) and (ii); either (\sim i₁) the twitch occurs after Jones's choice, and thus Black is too late to cut off alternate possibilities, (\sim i₂) the twitch occurs before Jones's choice, but is not causally connected to his choice, and thus Black cannot be sure he's cut off alternate possibilities (despite Black's prior observations, Jones can choose to \sim A without twitching), or (\sim ii) the twitch is causally connected to Jones's choice, such that his choice is wholly causally determined either by the twitch, or by a preceding states of affairs, and thus Frankfurt can't help himself to the conclusion that Jones is uncontroversially morally responsible for his actions.

In "Rescuing Frankfurt-Style Cases", Alfred Mele and David Robb introduce a method for circumventing this problem:

Black wants Bob to steal Ann's car by time t . He believes Bob will choose to freely steal the car by that time, but just in case he implants a device in Bob's brain that will causally determine him to steal Ann's car at t unless Bob has already freely chosen to steal it. As it so happens, Bob freely chooses to steal the car on his own, and the device doesn't play a causal role.

However, Bob's modified deliberation process is such that were he to freely choose to steal the car at t , his free choice is the cause, but if he would be such that he freely chooses to not to steal the car at t , his free choice is bypassed by the device.

The key innovation in this case is that the device in question is said to activate at time t regardless of what Bob would choose. If Bob would freely choose not to steal the car, or has simply failed to make a choice, the device bypasses Bob's normal deliberation method and forces him to choose to steal

⁴ See Robert Kane (1985, 1996); David Widerker (1995).

the car. If, however, Bob would freely choose to steal the car, although the device activates, it has no effect, and Bob chooses to steal the car of his own free will. The primary problem with this version of the case is that it's not clear that a device could be constructed in such a way that it would be ineffective when Bob freely chooses what Black wants, but effective when Bob would choose otherwise. In a sense, Mele and Robb want it both ways – Bob acts freely when he does what Black wants, and is bypassed before his choice when he counterfactually would have acted otherwise. The reason why Mele and Robb's device is said to be able to do this is because it activates during Bob's deliberation process, and they offer a mechanistic account of Bob's deliberation process such that the device is the cause if and only if Bob would counterfactually have freely chosen otherwise.

Both Fischer and Derk Pereboom construct Frankfurt-style cases where similar devices play an important role⁵, but fail to adopt Mele and Robb's account of how the device works, thus leaving their position open to the questions about how the device knows when to intervene.

Rather than adopt Mele and Robb's mechanistic causal story of agency, Fischer turns to the work of David Hunt and Derk Pereboom, in which they contend that the relevant question is not whether the agent in a Frankfurt-style case has alternate possibilities, but whether she has robust alternate possibilities.⁶ Pereboom defines *robustness* in the following way:

For an alternative possibility to be relevant per se to explaining an agent's moral responsibility for an action it must satisfy the following characterization: she could have willed something other than what she actually willed such that she understood that by willing it she would thereby have been precluded from the moral responsibility she actually has for the action. (2008, 4–5)

Pereboom's account builds upon Otsuka's PAB – requiring a difference in blameworthiness between options. However, this account is deficient. To have robust alternate possibilities on this account, one needs to have the option to act in a manner in which one would lack moral responsibility of any kind – but this is absurd! Moral agents are morally responsible for their free choices, blameworthy for their vicious ones, and praiseworthy – or at least blameless – for others. This is to say that when a moral agent acts as a moral

⁵ See Fischer (2010), Pereboom (2008).

⁶ See Hunt (2000), Pereboom (2001).

agent, even when she chooses to act in a way in which is neither praiseworthy nor blameworthy (for example, choosing between two morally equivalent options), she is still morally responsible for her choice. Otsuka's PAB requires the ability to act in such a way in which one is entirely blameless, not entirely lacking moral responsibility. Such an agent may be praiseworthy (as with Smith¹) or responsible but neither praiseworthy nor blameworthy. In "On Robust Alternate Possibilities and the Tax Evasion Case," I propose an account of robust alternate possibilities that includes Pereboom's awareness criteria, but more closely resembles Otsuka's account:

An agent has a robust alternate possibility if and only if she could have acted in a manner in which she believes [or would believe] she would have been differently morally responsible (2015a, 103).

On my view, for each action one is morally responsible for, one has at least one robust alternate possibility in which, if one is blameworthy, one would be entirely morally blameless, and if one is praiseworthy, one would be entirely morally praiseless.⁷

In Frankfurt-style cases, we have good reason to think the agent in question has robust alternate possibilities. In Frankfurt's original case it makes sense to say that Jones believes he has multiple possibilities – that he can freely choose to A or \sim A. Suppose Jones believes freely choosing A is *prima facie* more morally blameworthy than freely choosing \sim A. If Jones were to freely choose A, sans Black's device, his choosing \sim A would uncontroversially be a robust alternate possibility. But suppose Jones chooses \sim A believing it to be morally superior to A, and Black intervenes after this choice forcing him to A against his will. His free choice to \sim A is praiseworthy, if only because Jones believes it would lead him to \sim Aing; the fact that Black's machinations prevent him from doing what he chooses is irrelevant. When Jones As in such a case, he does so against his will and it doesn't make sense to hold him either blameworthy or praiseworthy as his actions are outside of his control.

If Black intervenes *before* Jones's choice, then Jones is not responsible because he doesn't choose anything; Black has bypassed his free will. But let us suppose Black uses a device like the one imagined by Mele and Robb – a device that intervenes *during* the choice, one that changes the freedom of his

⁷ When one acts in a blameworthy fashion, one's robust alternate possibility might be such that she would be not only entirely morally blameless, but also such that she would be praiseworthy.

choice depended on the expected outcome – if Jones would freely A, the device does nothing and Jones’s free choice causes his Aing, but if Jones would counterfactually have freely \sim Aed, the device causally determined his Aing. Here too, I think, Jones has robust alternate possibilities. By assumption, Jones believes:

- (1) He would be blameworthy for freely choosing to A.
- (2) He would be differently responsible (less blameworthy, both blameless and praiseless, or praiseworthy) for freely choosing to \sim A.

However, surely Jones also believes:

- (3) He would lack *any responsibility at all* for the actions causally necessitated by a device that he didn’t willingly, knowingly, or negligently trigger.

Belief (3) is not a convenient, *ad hoc* belief – it seems to follow from the commonsense belief that moral responsibility requires control. In Mele and Robb’s case, if the device activates Jones would be causally responsible (in some sense) for its activation (as if he had just did as the device would have caused him to do, the device would have remained dormant); however he would have no more moral responsibility for this than a hiker that unintentionally and non-negligently steps in a well-disguised bear-trap along her normal jogging route (well known for its otherwise consistent lack of bear-traps – disguised or otherwise). This is to say that Jones’s causal responsibility would be insufficient to hold him morally responsible for the outcome.

Fischer admits traditional Frankfurt cases fail to show that the agent lacks robust alternate possibilities, but he contends that “buffer zone” cases, like ones proposed by Pereboom and Hunt, are promising.⁸ In buffer zone cases, it is stipulated that before one is psychologically allowed to consider doing otherwise, they have to meet certain criteria, but some device either prevents them from fully meeting these criteria, or prevents them from acting freely once the criteria is met. A concise version of Pereboom’s case goes something like this:

Tax Evasion. Joe believes he can get away with cheating on his taxes, but that to do so would be wrong. Joe’s strong desire to advance his own self-interest will causally determine him to cheat on his taxes at t_1 unless he chooses against it for moral reasons. To choose against it for moral reasons, Joe must first raise his moral attentiveness to an

⁸ See Pereboom (2001).

appropriate level through the exercise of his libertarian free will. (Joe cannot act on a whim.) However, even if he were to attain this level of moral attentiveness, his libertarian free will allows him to choose either to pay his taxes or cheat.

Unbeknownst to Joe, a neuroscientist has implanted a device in his brain which is triggered by him reaching the appropriate level of moral attentiveness. If triggered, the device would rob him of his libertarian free will and causally determines him to cheat on his taxes. As it so happens, Joe never uses his libertarian free will to raise his moral attentiveness level “and he chooses to evade taxes on his own, while the device remains idle.” (2001, 9–10)

Pereboom contends two things are true in this case:

- (i*) Joe (uncontroversially) lacks robust alternate possibilities.
- (ii) Joe is uncontroversially morally responsible.

Perhaps the most notable problem with buffer zone cases is that they draw our attention away from the relevant issue – Joe’s free choice. To explain how the device can successfully cut off (robust) alternate possibilities, Pereboom has to stipulate that Joe’s deliberation process works in a rather specific way – a way that we’re told is libertarian/indeterministic, despite the fact Pereboom expressly contends it does not allow for Joe to act on a whim (the ability to act on a whim would constitute a robust alternate possibility, so Joe can’t possess it.) If Joe were to act in such a way that, *sans* device, would allow him to choose not to cheat on his taxes, the device would activate, interfering with, and bypassing, Joe’s normal deliberation method in such a way that Nahmias would conclude is responsibility-undermining. But don’t fret, we’re told that as it so happens Joe never bothers to try to do otherwise, he does what he was predisposed to do, and the device does nothing. Thus, Pereboom contends, it makes sense to say Joe freely chooses to cheat on his taxes – he does what he wants to do – and questions about the plausibility of the Joe’s deliberation process or the device fall by the wayside, as they don’t play a role in his choice.

This is quite a feat of sleight of hand – Joe’s choice architecture is said to be indeterministic, despite his inability to act on a whim. Joe’s said to have alternate possibilities, but due to the device these alternate possibilities can never change how he will behave. But as it so happens, Joe freely chooses not to act on these (ineffective) alternate possibilities, and he does what he was pre-disposed to do, and despite not exercising his indeterministic free

will, we're told this is his free choice. One could imagine Pereboom saying "Pay no attention to the choice architecture behind the curtain" that is said to cut off robust alternate possibilities, but all this serves to obfuscate the feature of the case that is actually causally effective – how Joe comes to be predisposed to act in his self-interest in the first place.

Suppose that a while back, Black wanted Joe to act in his self-interest and implanted a separate device that bypasses his normal deliberation process and forced Joe to be predisposed to act in his self-interest. If Joe's current predisposition to act self-interestedly is outside of his control in this way, both compatibilists and incompatibilists alike would conclude Joe is blameless for this predisposition. The fact that his (supposedly) libertarian free will would (*sans* neuroscientist's device) give him a second chance to fix this mistake is only somewhat relevant to determining his freedom or responsibility for acting on this predisposition, but in Tax Evasion, Joe would lack even this second chance.

If Joe's predisposition to act self-interestedly was brought about deterministically, incompatibilists would reject (ii). However, if it was brought about in an appropriate indeterministic way, then Pereboom has failed to show that (i*) is true, as Joe might very well have had robust alternate possibilities with regards to setting his predisposition – for example, either he could freely choose to act in such a way that would lead to a predisposition to act self-interestedly or to act in some other way. Complicating the issue, acting in such a way that conditions one to act self-interestedly is not inherently immoral – often the right thing is also the self-interested thing to do – for example, it is often both in your self-interest and morally correct to exercise, eat right, and practice good hygiene. Thus, even if Joe freely acted in such a way that conditioned himself to act in his self-interest, this is not necessarily morally blameworthy.

But now, let us look at a feature of Joe's deliberation process necessary, but not sufficient, to get him to reconsider cheating on his taxes – raising his moral attentiveness level. Assume the following:

- 1) Joe knows that he can only change his predisposition by raising his moral attentiveness level.
- 2) Joe knows how to raise his moral attentiveness level.
- 3) Joe believes that if he does nothing, he will act immorally by cheating on his taxes.
- 4) Joe believes there are no other, more pressing morally relevant concerns that ought to take precedence.

If these are true, then I think it makes sense to say that Joe *does* have robust alternate possibilities – he can either freely choose to raise his moral attentiveness level, or he can freely neglect to do so. Consider the following principle:

Stepping Stone Principle (SSP): If *a* has a moral obligation to do *x*, and *a* believes *y* is necessary to do *x*, then *a* has a moral obligation to do *y*.

Joe believes that a necessary (although not sufficient) step to avoid cheating on his taxes is to raise his moral attentiveness level, and if SSP or something like it is true, Joe has a moral obligation to raise his moral attentiveness level so that he can avoid cheating on his taxes. Failure to do so is *prima facie* blameworthy. This means that Joe does have robust alternate possibilities in Tax Evasion – he can either raise his moral attentiveness level or he can refrain from doing so. He believes he ought to do the former, yet does the latter, and thus he fails to do what he believes he ought to do. But this *just is* to say that he had robust alternate possibilities – possibilities that, if he acted on them, would lead to his being differently responsible.

In contrast, if we do not assume 1–4, Joe is quite the bizarre agent, inexplicably ignorant about how to make decisions for himself. Such an agent would be incompetent and it wouldn't make sense to hold him blameworthy for cheating on his taxes because he lacked the awareness to do otherwise. At the very least, Joe would not be *uncontroversially* morally responsible for his actions.

Fischer admits he fails to offer a satisfactory defense against the indeterministic horn of the dilemma, but says:

I simply want to motivate the idea that the jury is still out with respect to the indeterministic horn of the Dilemma Defense. I wish to emphasize that it is not enough to point out that if indeterminism obtains, there will always be *some sort of residual alternative possibility*; the alternative possibility must be of the right sort—it must be sufficiently robust to ground attributions of moral responsibility. Arguably, explicitly indeterministic versions of the Frankfurt cases can be developed in which it is highly plausible that the agent is morally responsible and yet lacks access to robust alternative possibilities. Intuitively, the lack of access to robust alternative

possibilities in these cases is *irrelevant* to the agent's status as morally responsible (2010, 323).

Fischer's discussion here isn't meant to be definitive. However, it is a substantial attempt to reframe the scope of Frankfurt-style cases; like Pereboom, Fischer wants to move the bar – a successful Frankfurt-style case, he contends here, need only aim to show merely that an agent (uncontroversially) lacks robust alternate possibilities, not just any kind of alternate possibilities.

If an indeterministic Frankfurt-style case can be developed in which an agent (i) uncontroversially lacks any sort of residual alternate possibilities, and yet (ii) is uncontroversially morally responsible, then one would have robbed PAP of its intuitive plausibility, weakening the appeal of incompatibilist stances on moral responsibility. The problem, it seems, is that critics of PAP have not been able to construct such a case.

In absence of such a case, Fischer contends that it should be possible to develop an indeterministic Frankfurt-style case in which an agent (i*) (uncontroversially) lacks robust alternate possibilities, yet (ii) is uncontroversially morally responsible. There are two problems with this approach: (1) It seems critics of PAP have thus far been unable to construct cases of this kind either, and (2) even if critics of PAP are right and any old alternate possibilities are irrelevant to determining an agent's moral responsibility, it is still possible that they correspond with moral agency. Elsewhere, I argue that even residual alternate possibilities may be important indicators of moral responsibility (2015b). Both sides of the free will debate contend moral responsibility requires control, and (at least for incompatibilists), the lack of alternate possibilities is a rather noteworthy indicator that an agent lacks control even if their absence is not the cause of the agent's lack of control.

B. The Deterministic Horn

In an earlier paper "Recent Work on Moral Responsibility," Fischer presented an argument that the deterministic horn of the dilemma defense fails; and offers a refined version of that argument in "The Frankfurt Cases: The Moral of the Stories." He says:

[S]upposing that we explicitly assume that causal determinism obtains in the cases, it is important first to note that I do not propose that we precipitously conclude, from mere reflection on the cases, that (say)

Jones is morally responsible for his choice and action. Rather, the initial conclusion is that if he is not morally responsible, it is not because he lacks appropriate alternative possibilities. *This initial conclusion does not beg the question against the incompatibilist* (2010, 323).

Fischer proposes that the compatibilist take this intuition seriously, and examine all the other (non-PAP) reasons that one might conclude that Jones lacks moral responsibility in Frankfurt's case. He continues:

If such a theorist concludes that, since there are no *other* reasons that constitute good and sufficient reasons to believe that causal determinism rules out moral responsibility, causal determinism is indeed compatible with moral responsibility, this too would *not* beg the question against the incompatibilist (2010, 324).

The primary problem with Fischer's approach is that it doesn't preserve the open-endedness of Frankfurt's argument. The reason PAP has played such a dominant role in the discussion of free will problem is that it allows philosophers to sidestep the complex metaphysical and metaethical disputes at the heart of the free will debate. Frankfurt says that "Practically no one... seems inclined to deny or even question that [PAP] (construed in some way or another) is true" (1969, 829). This is to say that prior to Frankfurt's analysis, compatibilists and incompatibilists largely agreed that PAP (or something like it) was true. This presented a clear problem for the compatibilist position – it was quite an uphill battle to show that alternate possibilities are compatible with universal causal determinism, the theory that there is only one possible future.

The incompatibilists argument is clear and persuasive:

1. PAP – Moral responsibility requires alternate possibilities.
2. If determinism is true, there are no alternate possibilities.
3. Conclusion: Moral responsibility is incompatible with determinism.

I believe the primary virtue of Frankfurt's argument against PAP is that it, too, is intended to circumvent complex metaphysical and metaethical disputes at the heart of the free will debate:

1. Jones (i) uncontroversially lacks alternate possibilities and (ii) is uncontroversially morally responsible.

2. If PAP, either Jones (\sim i) has alternate possibilities or (\sim ii) is not morally responsible.
3. Conclusion: PAP is false.

Frankfurt's argument is not an argument for compatibilism, but rather an argument against PAP – if successful, all Frankfurt will have shown is that PAP is false and the (apparent) inconsistency between determinism and alternate possibilities is irrelevant because alternate possibilities are not required for moral responsibility.

The primary virtue of this argument is that it's open-ended – if you're predisposed to believe (i) and (ii), then you are predisposed to believe PAP is false. In contrast, Fischer argues that if you are predisposed to believe (\sim ii), this isn't enough to show the truth of PAP. But this misses the point – the incompatibilist appealed to PAP simply because most people (compatibilist and incompatibilist, philosopher and layman alike) already believe PAP (or something like it) is true. A successful Frankfurt-style case would immediately undermine the intuitive plausibility of the moral principle that has allowed incompatibilists to claim their position is *obviously* more consistent with our intuitions than the compatibilist position. Thus, without PAP to bolster the intuitiveness of their position, it is an open question which theory – compatibilism or incompatibilism – is more consistent with our moral intuitions. Of course, compatibilists believe they have a slight edge – our leading scientific theories assume our world acts deterministically, and so the compatibilist position would allow those who assume universal causal determinism to allow for moral responsibility, whereas incompatibilism would lead those who assume universal causal determinism to conclude no one is morally responsible for anything.

Fischer contends that even if you assume determinism and conclude that Jones is not responsible, either (a) you will find there are other responsibility-undermining features of the case independent of the assumption of determinism, or (b) you're begging the question and just assuming determinism itself undermines moral responsibility. Not so.

Regarding (a); compatibilists believe that their view is consistent with most of our commonsense moral intuitions, and thus they believe that upon closer reflection, there is a good chance that one will find other responsibility-undermining features of the case. But this is not the win Fischer thinks it is, as the very fact that arriving at this conclusion requires reflection demonstrates they failed to show (ii). The lack of a Frankfurt-style case that can show both (i) or (i*) and (ii) is quite a problem for the

compatibilist, as if determinism and compatibilism are true, then every instance of genuine moral responsibility is a counter-example to PAP. Yet faced with the threat of the intuitive plausibility of PAP, compatibilists have failed to construct a successful counter-example to PAP. Given the motivation of compatibilists, and their contention that such counter-examples are plentiful, their inability to point to an example that can do both (i) and (ii) is quite foreboding.

Note that concluding (a) is not altogether unproblematic for the proponent of Frankfurt-style cases. If there is some *other* responsibility-undermining feature of a Frankfurt-style case, then it's not a successful Frankfurt-style case (and thus fails to be a counter-example to PAP), but worse still it means that the architect of the case has an unreliable moral intuition – they believe the agent is morally responsible when, in fact, there are responsibility-undermining features of the case that they have overlooked. Not only would such a case fail to show (ii), it would fail to show even (ii*) that the agent is (controversially) morally responsible.

However, Fischer's contention that (b) those that do not find Frankfurt-style cases compelling are question-begging, is more problematic. First Fischer suggests that it is question-begging to conclude that a lack of alternate possibilities (brought about by determinism) is responsibility-undermining. However, before Frankfurt's article, both compatibilists and incompatibilists alike found PAP intuitively plausible and offered distinct accounts of what it required, with incompatibilists arguing PAP required actual alternate possible futures, and compatibilists arguing PAP required something like the ability to do otherwise should one have chosen to – counterfactual alternate possibilities of one kind or another. Thus, if one had the intuition that the agent in a Frankfurt-style case isn't morally responsible, this isn't to say that they're merely begging the question that determinism is incompatible with moral responsibility. After all, many compatibilists – including most that came before Frankfurt – would be happy to explain an agent's lack of moral responsibility in some cases on their lack of (compatibilist) alternate possibilities. For example, Nahmias can sensibly say that when an agent's normal deliberation process is bypassed, say by a neuroscientist's device, she lacks the ability to do otherwise if she had wanted to.

However, there is a deeper problem with Fischer's approach here – it gives up on the open-endedness of Frankfurt's original case. Frankfurt believed that everyone – compatibilist and incompatibilist, philosopher and

layman alike would conclude Jones (i) uncontroversially lacks alternate possibilities, and yet (ii) is uncontroversially morally responsible, undermining the intuitive plausibility of PAP and any intuitive plausibility the incompatibilist position gained by piggybacking on the principle. In contrast, Fischer's contention is that Frankfurt-style cases are plentiful, but that if you doubt the agent in any given Frankfurt-style case is (ii) uncontroversially morally responsible, then (a) don't worry – although it's not a genuine Frankfurt-style case, there is something other than a lack of alternate possibilities undermining moral responsibility, or (b) your intuition is unreliable. While Frankfurt's argument turns on our intuitions, Fischer's argument seems to do the opposite – relying upon intuitions only when they're consistent with compatibilism – in (a) – and dismissing them when they're not – in (b). Thus, it seems Fischer isn't only giving up (i), but (ii); for Fischer, all a counter-example to PAP would have to show is:

An agent:

(i*) lacks robust alternate possibilities.

(ii*) is actually (but possibly controversially) morally responsible (regardless of whether you have this intuition or not).

If Frankfurt is successful, the debate over PAP is over. However, if a Fischer-style Frankfurt-style case could be constructed to show (i*) and (ii*), debate would continue because there is no consensus on what alternate possibilities matter and whether the agent is genuinely morally responsible. This is to say that even if Fischer is right and compatibilism is true, he will have failed to demonstrate this in a persuasive way.

III. Conclusion

Here, I've argued that Frankfurt-style cases are an elegant attempt to reframe the free will debate, circumventing complex philosophical questions about metaethics and metaphysics and resting solely on our commonsense intuitions. If Frankfurt is right, then anyone who reads Jones's case would conclude that he (i) uncontroversially lacks alternate possibilities, and yet (ii) is uncontroversially morally responsible for his actions. Such a case would be a counter-example to PAP, and undermine its intuitive plausibility, in turn undermining the intuitive plausibility of a popular incompatibilist argument against compatibilism.

The primary problem facing proponents of Frankfurt-style cases is that – at least thus far – compatibilists have been unable to construct a case in which they can help themselves to both (i) and (ii). The dilemma defense

argues that either (\sim i) or (\sim ii). If Jones's actions are undetermined, then either Black would intervene before Jones's choice, undermining his agency and responsibility such that (\sim ii), or Black would intervene after Jones's choice, so Jones could have chosen otherwise such that (\sim i). However, if Jones's actions were determined, anyone that believes determinism undermines moral responsibility would conclude (\sim ii).

Fischer takes issue with the deterministic horn of the dilemma, contending that it's not fair to conclude that Jones lacks responsibility just because the case assumes determinism. He contends that there may be other reasons why Jones lacks responsibility, or we ought to conclude he is responsible. This is absurd. Our moral intuitions regarding freedom and responsibility are nuanced and complex – the virtue of Frankfurt's strategy is that he intends to circumvent any serious discussion about them by appealing to a case in which he believes everyone – regardless of their moral and metaphysical intuitions – will conclude the agent is not only responsible, but uncontroversially so. If everyone agrees Frankfurt's case is a counter-example to PAP, the debate is over.

As it so happens, Frankfurt's case is not as uncontroversial as intended. Although he intended the case to be metaphysically neutral, the dilemma defense illustrates that he has failed to do so. Fischer contends explicitly indeterministic Frankfurt-style cases could be developed, but refrains from doing so. This is disappointing, as such a case would be sufficient to convince the target audience (incompatibilists) that PAP is unintuitive.

Instead, Fischer focuses on the deterministic horn, challenging the intuitions of those that are unsure of an agent's responsibility in a deterministic Frankfurt-style case. But challenging intuitions in such a case undermines the open-ended, prereflective nature of Frankfurt's argument – in the same way compatibilist Fischer questions the validity of intuitions inconsistent with his metaethical stance on moral responsibility, so, too, could incompatibilists – those that believe responsibility is incompatible with determinism – and hard incompatibilists – those that believe responsibility is incompatible with either determinism or indeterminism – challenge the intuitions of those that agree with Fischer. In other words, rather than circumvent questions about our metaphysical and metaethical beliefs, as Frankfurt and those that appeal to PAP do, Fischer bogs the discussion down in a quagmire of complexity and skepticism regarding (at least some of) our moral intuitions. Perhaps this is as it needs to be – our intuitions regarding metaethics and metaphysics are nuanced and complex and require rigorous

analysis to resolve important questions regarding freedom and responsibility. However, the appeal of Frankfurt-style cases is that they are meant to skip all that.

The literature surrounding Frankfurt-style cases have helped ethicists to make substantial advances in clarifying their position on free will and moral responsibility, and made it easier to articulate important normative principles distinct from PAP, such as Michael Otsuka's principle of avoidable blame. Both compatibilists and incompatibilists contend that their account of moral responsibility is consistent with our moral intuitions. Incompatibilists, it seems, have the upper hand in this debate, owing to the intuitive plausibility of PAP. Frankfurt-style cases represent a means by which compatibilists can challenge the intuitive plausibility of PAP.

If determinism is true, and moral responsibility is compatible with determinism, one would expect successful deterministic Frankfurt-style cases to be quite common and easy to construct. Constructing even a successful indeterministic Frankfurt-style case would be sufficient to undermine the intuitive plausibility of PAP. Yet compatibilists have thus far failed to construct cases of either kind. If our intuitions are compatibilist, the lack of such cases is shocking. However, if our intuitions are incompatibilist, this failure should come as no surprise.

Attempts by Pereboom, Hunt, and Fischer to reframe the debate by focusing on robust alternate possibilities and/or questioning the intuitions of those that don't find the agent in question uncontroversially morally responsible seem defeatist. At best, this muddles the debate in complex philosophical issues, rather than circumvent those issues entirely as Frankfurt intends. Incompatibilists appeal to PAP to circumvent those very issues; if compatibilists hope to show their view is as intuitively plausible as the incompatibilists, they should keep this in mind and, like Frankfurt, attempt to follow suit.

References

- Bernard Berofsky, (2012). *Nature's Challenge to Free Will*. Oxford University Press.
- Fischer, John Martin and Ravizza, Mark (1998). Morally Responsible People without a Freedom. from *Responsibility and Control: A Theory of Moral Responsibility*. New York: Cambridge, from *The Determinism and Freedom Philosophy Website*, accessed: 7-05-09
<http://www.ucl.ac.uk/~uctytho/dfwCompatFischerRavizza.htm>

- Fischer, John Martin (1999). Recent Work on Moral Responsibility. *Ethics* 110(1), pp. 93–139.
- Fischer, John Martin. (2010). The Frankfurt Cases: The Moral of the Stories. *Philosophical Review* 119(3), pp. 315–336.
- Frankfurt, Harry G. (1969). Alternate Possibilities and Moral Responsibility. *The Journal of Philosophy* 66(23), pp. 829–839.
- Ginet, Carl. (1996). In Defense of the Principle of Alternative Possibilities: Why I Don't Find Frankfurt's Argument Convincing. *Philosophical Perspectives* 10, pp. 403–417.
- Goetz, Stewart (2005). Frankfurt-Style Counter-examples and Begging the Question. *Midwest Studies in Philosophy* 29, pp. 83–105.
- Hunt, David (2000). Moral Responsibility and Unavoidable Action. *Philosophical Studies* 97, pp. 195–227.
- Kane, Robert (1985). *Free Will and Values*. Albany: State University of New York Press.
- Kane, Robert (1996). *The Significance of Free Will*. New York: Oxford University Press.
- Kane, Robert (2003). Free Will: New Directions for an Ancient Problem. in Kane (ed.): *Free Will*, Blackwell.
- Mele, Alfred R. and Robb, David (1998). Rescuing Frankfurt-Style Cases. *The Philosophical Review*, 107(1), pp. 97–112.
- Nahmias, Eddy (2011). “Intuitions about Free Will, Determinism, and Bypassing”. in R. Kane, (ed.) *The Oxford Handbook of Free Will 2nd Ed.* Oxford University Press: 555–587
- Otsuka, Michael (1998). Incompatibilism and the Avoidability of Blame. *Ethics* 108(4), pp. 685–701.
- Pereboom, Derk (2001). *Living Without Free Will*, Cambridge: Cambridge University Press.
- Pereboom, Derk (2008). Defending Hard Determinism Again. in *Essays on Free Will and Moral Responsibility*, ed. Nick Trakakis and Daniel Cohen, Newcastle-upon-Tyne: Cambridge Scholars Press, pp. 1–33.
- Rachels, James (2003). *The Elements of Moral Philosophy Fourth Edition*, McGraw-Hill.
- Simkulet, William (2015a). On Robust Alternate Possibilities and the Tax Evasion Case. *Southwest Philosophy Review* 31(1), pp. 101–107.
- Simkulet, William (2015b). On the Signpost Principle of Alternate Possibilities: Why Contemporary Frankfurt-Style Cases are Irrelevant to the Free Will Debate. *Filosofiska Notiser* 2(3), pp 107–120.

William Simkulet

Strawson, Galen (1994/2002) The Impossibility of Moral Responsibility. *Philosophical Studies* 75, pp. 5–24. reprinted in *Ethical Theory Classic and Contemporary Readings* by Louis P. Pojman.

Widerker, David (1995). Libertarianism and Frankfurt's Attack on the Principle of Alternative Possibilities. *Philosophical Review* 104, pp. 247–261.

William Simkulet
Cleveland State University
simkuletwm@yahoo.com