

Alternative possibilities in context

Alex Kaiserman

Abstract: Frankfurt cases are often presented as counterexamples to the principle that one is morally responsible for one's action only if one could have acted otherwise. But 'could have acted otherwise' is context-sensitive; it's therefore open to a proponent of this principle to reply that although there is a salient sense in which agents in Frankfurt-style cases couldn't have acted otherwise, there's another, different sense in which they could have, and it is this latter sense which is relevant to what we are morally responsible for doing. In this paper, I will evaluate the prospects of this contextualist response. I will argue that despite some initial signs of promise, the response fails, for reasons that were clearly anticipated in Frankfurt's original paper.

Key words: Frankfurt-style cases; contextualism; alternative possibilities; moral responsibility; free will

1. Introduction

Frankfurt's seminal 1969 article is best known as the source of a family of purported counterexamples to the *principle of alternate possibilities*:

PAP: S is morally responsible for her action only if she could have acted otherwise.

Here's a familiar example of a 'Frankfurt-style case':

Black and Jones: Black, a neuroscientist and committed Democrat, has been secretly monitoring the brain processes of Jones, who is deliberating over whether to vote Democrat or Republican. Black can reliably predict the choices that Jones is about to make by looking at the activity in his brain, and can also manipulate Jones's brain in a way that guarantees that Jones will vote Democrat. He plans to intervene if he predicts that Jones will not vote Democrat on his own. As it happens, Jones chooses to vote Democrat on his own, motivated by his own reasons, and without the intervention of Black (who correctly predicts that Jones would vote Democrat on his own).

There's a clear sense in which Jones couldn't have acted otherwise in **Black and Jones**. Yet intuitively, he is morally responsible for voting Democrat. So the PAP is false, or so it is argued.

Frankfurt-style cases have, of course, generated a bewildering variety of responses. Some insist that while Jones couldn't have refrained from voting Democrat in **Black and Jones**, there are nevertheless various things he *could* have done – he could have *tried* not to vote Republican, for example – and it is in virtue of these facts that Jones is morally responsible for what he did.¹ Others insist that while Jones couldn't have acted *otherwise*, he could nevertheless have failed to act as he did, because had Black intervened, Jones's bodily movements wouldn't have constituted an *action*, strictly speaking, at all – and it is in virtue of *this* fact that Jones is morally responsible for voting Democrat.²

But there is a different way of responding to Frankfurt-style cases in defence of the PAP, one which I want to explore in this paper. The response begins with the observation that 'could have done otherwise', like most modal locutions, is *context-sensitive* – its meaning depends, roughly speaking, on which facts the context invites us to hold fixed. It is thus open to a defender of the PAP to insist that, although there is a salient sense in which Jones couldn't have acted otherwise in **Black and Jones**, one where we hold fixed, in particular, facts about Black, there is also a sense in which he *could* have acted otherwise, one where these facts are not held fixed; and it is this *latter* sense of 'could have done otherwise' relative to which the PAP ought to be interpreted.

The goal of this paper is to evaluate the prospects of this contextualist response in more detail. I will argue that despite some initial signs of promise, the response fails, for reasons that were clearly anticipated in Frankfurt's original paper. Though it might be possible to articulate a sense of 'could have done otherwise' that reconciles the PAP with our intuitions about Frankfurt-style cases, this is achievable only by helping oneself to the very facts which according to Frankfurt are what really *ground* facts about moral responsibility in the first place – namely, facts about what caused the agent's action. I conclude that even if there are true, non-

¹This is often referred to as the 'flicker defence' by its detractors, who typically argue that such alternatives are not 'robust' enough to ground moral responsibility for one's actions. For discussion, see Fischer (1994, ch. 7) and Mele (1998).

²See, e.g., Alvarez (2009).

trivial interpretations of the PAP, it doesn't follow that their truth tells us anything about what it *is* to be morally responsible for one's actions.

2. Contextualising the PAP

It's well-known that modal claims – claims about what may, might, can, could, must, has to or ought to be the case – can be used to mean very different things in different contexts of use. The sentence 'Tom can't speak French', for example, might be a reminder of the rules (Tom's not allowed to speak French), an expression of concern (it's not in Tom's best interests to speak French), a claim about Tom's cognitive skills (Tom is not a French speaker) or physiological capacities (Tom has had his larynx removed), or a claim about the environmental conditions (Tom is gagged). Yet according to the orthodox semantics of these expressions,³ there is nevertheless a core logical structure to words like 'can' and 'must' that is invariant across their various uses. An utterance of the sentence 'Tom can speak French', on this view, expresses a proposition which is true if and only if Tom speaking French is metaphysically compossible with some contextually-determined set of facts (the facts we're 'holding fixed' in the context). Thus we get different interpretations of 'Tom can't speak French' depending on whether the context invites us to hold fixed the fact that no rules are broken, facts about Tom's physiological or psychological properties, facts about the environmental conditions, and so on.

It follows immediately, if this view is correct, that sentences like 'Jones could have acted otherwise' are context-sensitive – their meaning depends on which facts the context invites us to hold fixed. There are contexts relative to which 'Jones could have acted otherwise' is trivially true, ones in which *nothing* is held fixed;⁴ contexts relative to which 'Jones could have acted otherwise' is trivially false, ones in which the fact that Jones *didn't* act otherwise is held fixed;⁵ and contexts relative to which 'Jones could have acted otherwise' is true only if determinism is false, ones in which the laws of nature and state of the universe in the distant past are held fixed. But there are also all kinds of other contexts in between these extremes. Suppose I drank

³See especially Kratzer (1977); c.f. Lewis (1976).

⁴Assuming, of course, that it is not essential to Jones that he perform the act in question.

⁵As Lewis (1976) points out, it's easy to find oneself in such 'fatalistic' contexts when considering whether time-travellers can kill their past selves.

coffee this morning. Could I have acted otherwise? Well, that depends; one might be asking whether I could have acted otherwise in view of certain facts about my psychological dispositions (am I addicted to coffee?), facts about the external conditions (was there anything else in the house to drink?), facts about my or others' wellbeing (was someone threatening to kill me if I didn't drink coffee?), and so on. We can gesture at what we mean explicitly, but most of the time we needn't bother – the context is usually sufficient to fix the meaning for us.⁶

It might seem absurd to maintain that there is any sense in which I could have acted otherwise if my action was determined by the laws of nature and the state of the universe in the distant past. But the contextualist has a way of explaining away this sense of absurdity. To see this, it's useful to consider an analogy with contextualism about ascriptions of knowledge. Lewis begins his 'Elusive Knowledge' with the following exposition of the problem of scepticism:

It is a Moorean fact that we know a lot. It is one of those things that we know better than we know the premises of any philosophical argument to the contrary. [But] no sooner do we engage in epistemology ... than we meet a compelling argument that we know next to nothing. [...] [I]t seems as if knowledge must be by definition infallible. If you claim that S knows that *P*, and yet you grant that S cannot eliminate a certain possibility in which *not-P*, it certainly seems as if you have granted that S does not after all know that *P*. [But] uneliminated possibilities of error are everywhere. [...] We never have infallible knowledge.

So we know a lot; knowledge must be infallible; yet we have fallible knowledge or none (or next to none). We are caught between the rock of fallibilism and the whirlpool of scepticism. Both are mad!

(Lewis 1996, 549–50)

Lewis's solution to this problem is to defend a contextualist semantics of knowledge ascriptions – an utterance of the form 'S knows that *p*' expresses a proposition which is true if and only if *not-p* is metaphysically impossible with S's evidence *together* with some contextually-

⁶'When we talk to each other, we hardly ever make explicit in view of which circumstances something should be necessary or possible. We may give hints. Usually people understand. And they all understand in pretty much the same way.' (Kratzer 1981, 53–4).

determined set of facts: the facts being ‘properly presupposed’ in the context.⁷ Many everyday ascriptions of knowledge are true, on this view – even though our evidence rarely, if ever, rules out all possibilities of error – because they’re made in contexts where the falsity of farfetched sceptical hypotheses is being properly presupposed. But knowledge is also ‘elusive’, according to Lewis, in the following sense: to even consider a sceptical hypothesis is often to find oneself in a context in which the falsity of this hypothesis can no longer be properly presupposed (and hence a context relative to which one knows next to nothing). So while there are contextual parameters relative to which ‘I know I’m not a brain in a vat’ is true, it’s difficult, if not impossible, to truly utter this sentence, since the very act of uttering it is usually sufficient to change the context to one in which it expresses something false – *that’s* why it seems like an absurd thing to say.

The contextualist about ‘could have done otherwise’ can say something similar in response to the incompatibilist. Many everyday assertions of the form ‘S could have done otherwise’ express propositions whose truth is perfectly compatible with determinism, on this view. But alternative possibilities in a deterministic world are ‘elusive’, in the following sense: to even consider the thesis of determinism is often to find oneself in a context relative to which the laws of nature and state of the universe in the distant past are held fixed (and hence one relative to which I could have acted otherwise only if determinism is false). So while there are contextual parameters relative to which ‘My action was determined, yet I could have acted otherwise’ is true, it’s difficult, if not impossible, to truly utter this sentence, since the very act of uttering it is often sufficient to change the context to one in which it expresses something false – *that’s* why it seems like an absurd thing to say.⁸

⁷Or equivalently: ‘S knows that *p*’ is true iff S’s evidence rules out all possibilities in which not-*p* – *psst!* – except those possibilities we are properly ignoring. The ‘*psst!*’ clause is necessarily *sotto voce*, for Lewis, because to attend to a possibility is necessarily to cease to ignore it, so *a fortiori* to cease to properly ignore it. See Blome-Tillmann (2009) for criticisms of Lewis’s ‘rule of attention’.

⁸Of course, the basic idea here is not particularly new. Kratzer herself hints at it in the following passage:

Suppose a judge asks himself whether a murderer could have acted otherwise than he eventually did ... Given the whole situation of the crime, which includes of course all the dispositions of the murderer, this man could not have acted otherwise than as he did ... [But t]he answer to the question of the judge is *not* trivial ... what the judge probably meant was: Given such and such aspects of the situation, could the murderer have acted otherwise than he eventually did? (Kratzer 1977, 343).

If ‘could have done otherwise’ is context-sensitive in the way described above, what should we make of the status of the PAP? Again, the analogy with contextualism about knowledge ascriptions is instructive here. After all, there are also seemingly tight conceptual connections between knowledge and moral responsibility. Consider the following principle, for example:

Knowledge Principle (KP): S is not morally responsible for ϕ -ing at t if she didn’t know, at any time prior to t , that she would ϕ at t (unless she is morally responsible for the fact that she didn’t know).

KP, or something like it, seems very plausible.⁹ If Jane didn’t know that her neighbour’s cat was sleeping under her tyre, and she is not responsible for her ignorance, then she is not responsible for running over the cat. But now suppose that ascriptions of knowledge are context-sensitive. Must a proponent of the KP thereby conclude that ascriptions of moral responsibility are *also* context-sensitive? I think not. A contextualist can agree that there is a particular, context-independent epistemic standard Jane must have met with respect to the proposition that the cat was under her tyre in order to be morally responsible for running it over; they would simply add that while that there are some contexts relative to which the word ‘know’ expresses that relation which is instantiated if and only if Jane meets this standard, there are also other contexts relative to which it expresses a different relation, one governed by different standards. On this view, the KP is true in some contexts and false in others. The interesting question, therefore, is not *whether* there is a context relative to which Jane is not morally responsible for running over the cat if she didn’t know (and isn’t responsible for not knowing) that it was under her tyre, but rather *which* context that is; what ‘kind’ of knowledge, in other words, is the kind that is relevant to what one is morally responsible for doing. Suppose Jane had plenty of evidence that the neighbour’s cat was under her tyre – she saw a cat-like figure as she approached her car, heard a cat-like mewl as she turned on the engine, ignored several shouted warnings from onlookers as she started reversing, and so on. Intuitively, Jane

Others who have noted the potential relevance of contextualism about ‘can’ claims to debates over free will and moral responsibility include Campbell (1997; 2005), Szabo and Knobe (2013), Franklin (2015), List (2014), Grzankowski (2014), Slater (2005) and Jeppsson (2016). Hawthorne (2001), Rieber (2006), Willaschek (2010) and Jaster (2020) all defend contextualist accounts of *free will* ascriptions too, though, curiously, only Jaster explicitly relates this alleged context-sensitivity of ‘free’ to the context-sensitivity of ‘can’. I will not take a stand in this paper on whether ascriptions of free will are context-sensitive.

⁹For some recent work on the epistemic condition on responsibility, see Robichaud and Wieland (2017).

is morally responsible for running over the cat (supposing the non-epistemic conditions on moral responsibility are satisfied). Of course, there are contexts relative to which Jane didn't know that a cat was under her tyre, since Jane's evidence doesn't rule out the possibility that she's hallucinating the whole thing. But this doesn't necessarily show that the KP should be abandoned, or that Jane isn't morally responsible after all – all it shows, as Slater (2005, 129) puts it, is that 'the "kind" of knowledge required for moral responsibility falls short of the requirements assumed by the skeptic'.

A contextualist about 'could have done otherwise' can say something similar about the PAP. 'Could have done otherwise' is context-sensitive; ascriptions of moral responsibility are not. Thus the PAP is true in some contexts and false in others. The interesting question, therefore, is not *whether* there's a context relative to which I am morally responsible for acting only if I could have acted otherwise, but rather *which* context that is; what 'kind' of possibility, in other words, is the kind that is relevant to what I am morally responsible for doing.¹⁰ On this view, Frankfurt cases acquire a quite different dialectical significance to the one they are typically assumed to have. It's true of course that there is a natural context relative to which Jones couldn't have acted otherwise in **Black and Jones**, one where we hold fixed, in particular, the presence of Black. But there are also contexts relative to which Jones *could* have acted otherwise, since there are of course possible worlds in which Jones decides to vote Republican and Black, for whatever reason, doesn't intervene. What Frankfurt cases show, then, is not that the PAP should be abandoned, or that Jones isn't responsible for voting Democrat after all – what they show is rather that the 'kind' of possibility relevant to moral responsibility is one where facts about agents like Black are not held fixed.

So, at any rate, it might be argued. In the next section, I will argue that this contextualist approach to reconciling the PAP with our intuitions in Frankfurt-style cases ultimately fails. The reasons for this, moreover, were clearly evident in Frankfurt's original paper, though they have always received the attention they deserve. As we'll see, Frankfurt was well aware of the context-sensitivity of 'could have done otherwise', and carefully insulated his argument against attempts to exploit it.

¹⁰c.f. Franklin (2015, 2104), who similarly characterises the debate over the PAP as 'a dispute about which abilities to do otherwise are necessary for moral responsibility, not whether the ability to do otherwise is necessary.'

3. Grounding and gerrymandering

Let's start by considering a different kind of case:

Ann and Cecilia: Ann suffers from kleptomania which manifests in a strong compulsive desire to steal. One day, Ann's kleptomania causes her to steal Cecilia's bag.

As Frankfurt (1969, 830) notes, it is cases like these that are generally used to motivate the PAP in the first place. Ann is not morally responsible for stealing Cecilia's bag, and it seems plausible, at least initially, to explain this in terms of the fact that Ann couldn't have acted otherwise. The challenge for a proponent of the contextualist strategy outlined above is therefore to specify an interpretation of the PAP which implies that Ann is not morally responsible for her action in **Ann and Cecilia**, but *doesn't* imply that Jones isn't morally responsible for his action in **Black and Jones**.

What interpretation of 'could have done otherwise' might deliver this result? Consider first the following proposal due to List (2014; 2019).¹¹ According to List, the kind of possibility relevant to S's moral responsibility for acting at *t* is possibility *holding fixed S's exact psychological state prior to t* (and the fundamental laws of nature). If Ann's action 'was the inevitable result of a ... psychological compulsion', then she could not have acted otherwise, on this interpretation, as required (List 2019, 107). But since 'the same psychological properties of an agent can be instantiated by different configurations of physical properties of the underlying organism', each of which may result in a different action being performed, List argues that 'the totality of facts at the psychological level up to a given time may leave more than one course of action open for an agent', and so 'there may be agential indeterminism, even in the presence of physical determinism' (List 2019, 91–2). Moreover, as well as reconciling the PAP with physical determinism, this interpretation also seems to reconcile it with our intuitions about **Black and Jones**: it's compossible with Jones having the exact psychological state he actually has prior to voting Democrat (and the fundamental laws) that he decides to

¹¹Daniel Kodsí and I discuss List's view in more detail in Kaiserman and Kodsí (forthcoming).

vote Republican, Black doesn't intervene, and Jones votes Republican. So Jones could have acted otherwise, in List's sense, notwithstanding the presence of Black.¹²

Unfortunately, there are several serious problems with List's proposal. The first is that, at least as it stands, List's proposed interpretation of the PAP fails to imply that Ann isn't morally responsible for stealing Cecilia's bag in **Ann and Cecilia**. After all, no action is ever the *inevitable* consequence of a psychological compulsion. It's perfectly compatible with Ann's exact psychological state at any time t prior to t' that she doesn't steal Cecilia's bag at t' , because it's perfectly compatible with Ann's exact psychological state at t that someone intervenes by administering a serum at some time between t and t' which instantly cures Ann of her kleptomania. The same goes for every psychological condition normally thought to deprive its sufferers of moral responsibility for their actions. As it stands, then, List's proposed interpretation simply *trivialises* the PAP, since it ensures that the PAP is satisfied in pretty much every case, including those, like **Ann and Cecilia**, used to motivate the principle in the first place.

At some points, List seems to suggest that certain facts about the agent's environment – such as, perhaps, the absence of any kleptomania-curing serums in the vicinity – should also be included among the facts to be held fixed when talking about what was possible for them.¹³ But there's an obvious difficulty with this strategy. Presumably, we want to hold fixed facts about the absence of kleptomania-curing serums in cases like **Ann and Cecilia**, but *not* facts about Black in **Black and Jones**; and it's not clear that there's any principled way of drawing this distinction.

In fact the problem is worse than this, because, as Cohen and Handfield (2007) point out, one can easily imagine Frankfurt-style cases in which the 'back-up' mechanism is internal to the agent, rather than part of her external environment. Consider the following variant on **Ann and Cecilia**, for example:

¹²Curiously, though, List doesn't actually endorse this response to Frankfurt's argument – he advocates a version of the flicker defence instead (see List 2019, 168–9).

¹³'A complete specification of all agents' psychological states *and their macroscopic environment* does not entail a complete specification of all the underlying microphysical details' (List 2019, 91; my emphasis).

Ann and Cecilia*: Ann decides to steal Cecilia’s bag. Had she decided not to steal it, this would have triggered her latent kleptomania, resulting in an irresistible compulsive desire that would have caused her to steal Cecilia’s bag anyway. But as it happened, Ann’s action was performed willingly and motivated entirely by her own reasons – the kleptomania in fact played no role at all in bringing about her action.

It seems to me that our intuitions about **Ann and Cecilia*** perfectly mirror our intuitions about **Black and Jones**: Ann is morally responsible for stealing Cecilia’s bag, notwithstanding her kleptomania, just as Jones is morally responsible for voting Democrat notwithstanding the presence of Black. So if the PAP is to capture this difference between **Ann and Cecilia** and **Ann and Cecilia***, we presumably need an interpretation of ‘could have done otherwise’ on which Ann’s kleptomania is held fixed in the former case but not the latter. Yet the only difference between the two cases seems to be in what *caused* Ann’s action – in **Ann and Cecilia** her action is caused by her psychological disorder, whereas in **Ann and Cecilia*** Ann’s own reasons to steal the bag *pre-empt* her kleptomania, which plays no role in bringing about the action.

This, of course, is precisely Frankfurt’s point. Frankfurt was well aware that ‘could have done otherwise’ can be interpreted in different ways. Yet importantly, he didn’t find it necessary ‘to consider in what sense the concept of “could have done otherwise” figures in the principle of alternative possibilities’ (Frankfurt 1969, 834).¹⁴ The reason for this is that his argument against the PAP doesn’t actually rely on any particular counterexample, like **Black and Jones**; instead, what he provides is a *recipe* for generating counterexamples, given *any* reasonable interpretation of ‘could have done otherwise’ that might be proposed. This point appears to have been lost in the morass of replies and counterreplies that followed Frankfurt’s original paper. It’s worth returning once more to the relevant passage:

Suppose someone – Black, let us say – wants Jones to perform a certain action. [...] If it does become clear that Jones is going to decide to do something else, Black takes effective steps to ensure that Jones decides to do, and that he does do, what he wants him to do. [...] What steps will Black take, if he believes he must take steps, in order to ensure that Jones decides and acts as he wishes? Anyone with a theory concerning what ‘could have done

¹⁴‘To discuss the [PAP] without analyzing [“could have done otherwise”] may well seem like an attempt at piracy. The reader should take notice that my Jolly Roger is now unfurled.’ (Frankfurt 1969, 835).

otherwise' means may answer this question for himself by describing whatever measures he would regard as sufficient to guarantee that, in the relevant sense, Jones cannot do otherwise. [...] *Given any conditions under which it will be maintained that Jones cannot do otherwise, in other words, let Black bring it about that those conditions prevail.* [...] Now suppose that Black never has to show his hand because Jones, for reasons of his own, decides to perform and does perform the very action Black wants him to perform. In that case, it seems clear, Jones will bear precisely the same moral responsibility for what he does as he would have borne if Black had not been ready to take steps to ensure that he do it. (Frankfurt 1969, 835–6; my emphasis, subscripts omitted)

Here's what I take to be the recipe Frankfurt is proposing. First, invite the proponent of the PAP to describe a case in which the PAP, as they are choosing to interpret it, correctly predicts that an agent is not morally responsible for their action. In other words, invite them to describe a case in which an agent is not morally responsible for their action and couldn't, in whatever sense is deemed to be relevant, have acted otherwise. This, as Frankfurt notes, will invariably be a case in which some set of circumstances – 'hypnotic suggestion' or 'some inner compulsion', for example – 'both bring it about that a person does something and make it impossible for him to avoid doing it' (Frankfurt 1969, 830). Frankfurt's crucial insight was to see that these factors can come apart. In the second part of the recipe, we take the case the proponent of the PAP has offered and amend it by stipulating that the agent's *own* reasons *pre-empt* the state of affairs which caused the agent's action in the original case. This, Frankfurt suggests, will still be a case in which the agent couldn't have acted otherwise, in whatever sense the proponent of the PAP has in mind; yet the agent will seem perfectly morally responsible for their action, contrary to what the PAP predicts. This is exactly the recipe we followed above – start with a case, **Ann and Cecilia**, in which the PAP seems to get the right result; then construct an amended case, **Ann and Cecilia***, which is just like the first except that Ann's own reasons pre-empt her psychological disorder. Ann is morally responsible for her action in the second case but not the first. But, according to Frankfurt, any reasonable interpretation of 'could have done otherwise' will be one on which, if she couldn't have done otherwise in **Ann in Cecilia**, she couldn't have done otherwise in **Ann and Cecilia*** either.

Of course, the PAP is only a necessary condition on moral responsibility. It would be open to the proponent of the PAP therefore to claim that, in the relevant sense of 'could', Ann could have acted otherwise in *both* **Ann and Cecilia** and **Ann and Cecilia***, and what explains why

Ann isn't morally responsible in the former case is something else, a different necessary condition on moral responsibility. But this would be to underestimate the power of Frankfurt's recipe. Given *any* case in which it is alleged the PAP implies that an agent is not responsible for their action, Frankfurt's recipe purports to be able to construct a counterexample to the PAP from that case. Of course one could dodge this problem by choosing to interpret 'could' in the PAP in such a way that every agent could always have acted otherwise than they actually did, but this would be to strip the PAP of any content. Trivialising one's theory is too high a price to pay for avoiding counterexamples.

The only way for the contextualist to block Frankfurt's recipe, then, is by constructing an interpretation of 'could' on which whether a certain state of affairs caused an agent's action can make a difference to what the agent could have done. Though Frankfurt does not appear to have considered this option, it has been explicitly defended in the literature, most notably by Campbell (1997). According to Campbell, an agent could have acted otherwise, in the sense of relevance to the PAP, just in case there is an 'accessible world' in which they act otherwise, where accessible worlds are defined as follows:

A1. Accessible worlds must have the same laws of nature as the actual world.

A2. In accessible worlds, agents cannot have any abilities or capacities that they lack in the actual world.

A3. Any advantage that a person has in an accessible world which she lacks in the actual world must result from something that she does in that possible world.

A4. Accessible worlds need not include factors which are causally irrelevant to the performance of actions.

A3 is supposed to ensure that Ann *couldn't* have acted otherwise in **Ann and Cecilia** – in a world where Ann is administered a kleptomania-curing serum, for example, she has an 'advantage' she actually lacks which doesn't result from anything she does in that world, so such a world is ruled out. By contrast, A4 is supposed to ensure that Ann *could* have acted otherwise in **Ann and Cecilia*** – we can consider worlds in which Ann doesn't suffer from

kleptomania and chooses not to steal Cecilia's bag, but *only* because her kleptomania is causally irrelevant to the occurrence of her action.¹⁵

I don't doubt that, using something like Campbell's criteria, we might ultimately be able to articulate a sense of 'could have done otherwise' on which the PAP is neither false nor trivial. The question, however, is whether the truth of this principle would tell us anything about what it really *is* to be morally responsible for one's actions. Consider how Campbell's strategy looks to the person who thinks that what Frankfurt cases reveal is that facts about moral responsibility are partially grounded in facts about what *actually caused* the agent's action, rather than in facts about what the agent could or couldn't have done instead.¹⁶ From their perspective, Campbell's interpretation looks hopelessly gerrymandered, since it succeeds in reconciling the PAP with our intuitions about moral responsibility in Frankfurt cases only by helping itself to the very facts which, according to them, are what ground the facts about moral responsibility in the first place.

Here's an analogy to illustrate this point. Suppose I were to suggest the following necessary condition on being a bachelor: S is a bachelor only if S can walk on Mars. You might well respond by alleging that my principle is either false or trivial. If 'can' here is interpreted to mean *possible in light of current aerospace technology*, then no-one can yet walk on Mars, including every bachelor. But if 'can' is interpreted to mean simply *possible in light of Mars's gravitational field strength and typical human physiology*, then everyone can walk on Mars, so the principle is trivially satisfied. But I then protest that I've been misunderstood. The 'can' I had in mind (I explain patiently) was a special kind of possibility, one where the fact that S has not walked on Mars is held fixed if, but only if, S is not an unmarried man. I take it you would dismiss this move as nothing more than a fantastically misleading sleight of hand. What really

¹⁵Though Campbell is not completely clear on this point, I think we're supposed to interpret A4 as *overriding*, or qualifying, A2 and A3 – in accessible worlds, agents *can* have abilities, capacities or unearned advantages they don't actually have, *so long as* those abilities, capacities or advantages didn't actually play any causal role in bringing their action about. Otherwise, as Campbell appears to acknowledge, A3 would imply that Jones couldn't have acted otherwise in **Black and Jones** (because the worlds where Black isn't waiting ready to intervene are presumably ones where Jones has an 'advantage' he doesn't actually have which doesn't result from anything he does in that world).

¹⁶Something like this view is already implicit in Frankfurt's original paper, but for the most sophisticated contemporary development of it, see Sartorio (2016).

grounds facts about bachelorhood is not anything to do with the possibility of walking on Mars, it's the facts which went into constructing the gerrymandered interpretation of my proposed principle in the first place.

The challenge for those who, like Campbell, help themselves to facts about what caused the agent's action in articulating the relevant sense of 'could have done otherwise' is to explain why what they're doing is any more legitimate than my blatantly flawed attempt to salvage the Mars-walking condition on bachelorhood. The mere fact that it's possible to articulate an interpretation of 'could have done otherwise' on which the PAP is neither false nor trivial doesn't imply that the PAP, so-interpreted, tells us anything about what it *is* to be morally responsible for one's actions, since what really grounds moral responsibility might instead be the very facts used to articulate that interpretation in the first place. I think this point is often missed in the literature. Consider the following passage from Franklin (2015), for example:

According to Fischer, 'the moral of Frankfurt stories' is that they 'suggest that alternative possibilities are irrelevant to the grounding of moral responsibility' [quoting Fischer (2006, 198)]. At first glance, then, it seems that Fischer denies that an ability to do otherwise is necessary for moral responsibility. (Franklin 2015, 2096)

The second claim does not follow from the first. One can grant that there is a true (even *necessarily* true) interpretation of 'S is morally responsible for her action only if she could have acted otherwise', without having to think that what it *is* for S to be morally responsible for acting has anything to do with what she could have done instead, in any sense of 'could'. Something can be necessary for the truth of *p* without being part of what *makes it the case* that *p*. And what everyone in this debate is (or at least should be) interested in are the grounds of moral responsibility, not merely the truth or falsity of sentences of the form 'S is morally responsible for acting only if'.

With this distinction in view, what reason is there to think that it is the PAP, as Campbell is proposing to interpret it, which grounds facts about moral responsibility, rather than the causal facts which go into articulating his interpretation in the first place? Proponents of the PAP might point to the fact, noted above, that it seems intuitive to explain why Ann isn't morally responsible for stealing Cecilia's bag in **Ann and Cecilia** by citing the fact that she couldn't have acted otherwise. But we should be careful about appealing to intuitions here, because the interpretation of 'could have done otherwise' Campbell stipulates is not one which is obviously

available in natural language. To the extent to which it seems true to say ‘Ann couldn’t have acted otherwise’ about **Ann and Cecilia**, it seems equally true to say it about **Ann and Cecilia***. It’s not clear that we have any intuitions about the PAP as Campbell is actually proposing to interpret it.

Moreover, it’s worth recalling that Frankfurt does more than just offer a recipe for generating counterexamples to the PAP – he also offers an *error theory* of why it might seem natural to appeal to the impossibility of doing otherwise in explaining why someone wasn’t morally responsible for their action:

We often do, to be sure, excuse people for what they have done when they tell us (and we believe them) that they could not have done otherwise. But this is because we assume that what they tell us serves to explain why they did what they did. We take it for granted that they are not being disingenuous, as a person would be who cited as an excuse the fact that he could not have avoided doing what he did but who knew full well that it was not at all because of this that he did it. (Frankfurt 1969, 837–8)

If Ann pleads that she shouldn’t be held responsible for her action because, given her kleptomania, she couldn’t have acted otherwise, we are minded to accept her explanation, but only because we take it for granted that her kleptomania, as well as rendering it impossible for her to act otherwise, also *caused* her action. After all, it would have been disingenuous for her to cite as an excuse the fact that her kleptomania made it impossible for her to act otherwise, if she knew full well that it was not because of her kleptomania that she did what she did. If we were subsequently to learn that the kleptomania in fact played no role at all in bringing about Ann’s action, having been pre-empted by her own reasons, we would likely be much less inclined to pardon her. On the other hand, we *wouldn’t* be minded to excuse someone pleading that they shouldn’t be held responsible for their action because, given their values, reasons and principles, they couldn’t have acted otherwise – when Luther said, ‘Here I stand; I can do no other’, he wasn’t trying to avoid sanction, but on the contrary to take full responsibility for his actions.¹⁷ All this seems to suggest that moral responsibility for one’s actions has more to do with what caused one’s actions than with what one could or couldn’t, in any sense, have done instead. The latter is relevant, according to Frankfurt, only insofar as it is a guide to the former.

¹⁷See Dennett (1984, 133).

4. Conclusion

I began this paper by articulating a contextualist strategy for reconciling the PAP with our intuitions about Frankfurt cases. According to this strategy, what cases like **Black and Jones** reveal is not that the PAP is false, but rather that the sense of ‘could’ relative to which it ought to be interpreted is one where certain facts, like facts about Black, are not held fixed. In this way, one can grant both that Jones is morally responsible for his action in **Black and Jones**, *and* that there is a natural sense in which he couldn’t have acted otherwise than he actually did, while at the same time insisting that all this is compatible with the PAP, properly interpreted.

There is certainly an appeal to this strategy; indeed, I myself have expressed some sympathy for it in the past.¹⁸ Nevertheless, I now believe it fails. I don’t deny that ‘could have done otherwise’ is context-sensitive, and hence that there are interpretations of the PAP, ones which appeal to what caused the agent’s action in determining which facts to hold fixed, on which the PAP is neither false nor trivial. But it does not follow from this that what really *grounds* facts about moral responsibility is the truth of the PAP so-interpreted, rather than the causal facts that go into articulating the relevant interpretation of ‘could have done otherwise’ in the first place.

In making the case for this conclusion, I have drawn attention to two aspects of the argument Frankfurt advanced in his 1969 paper, which it seems to me have been relatively overlooked in the vast literature the argument has generated. The first is that Frankfurt’s argument does not rely on any particular counterexample to the PAP; rather, he provides a recipe for generating counterexamples, given any reasonable interpretation of ‘could have done otherwise’ that might be proposed. The second is that Frankfurt provides an error theory of the initial plausibility of the PAP, according to which we are willing to excuse people when they tell us that, given some set of circumstances, they couldn’t have acted otherwise, only because we take it for granted that these circumstances serve to explain why they did what they did. Together, these features of Frankfurt’s argument make it considerably more powerful than a

¹⁸See Kaiserman (2016). I arrived at the idea of contextualising the PAP after defending a similar approach to the idea that causes *necessitate* their effects, though I am yet to recant on the latter view (see Kaiserman (2017)).

particular purported counterexample, presented in isolation, would make it seem. There is more to Frankfurt's argument, in other words, than Frankfurt-style cases.¹⁹

References

- Alvarez, Maria. 2009. Actions, thought-experiments and the 'principle of alternate possibilities'. *Australasian Journal of Philosophy* 87: 61–81.
- Blome-Tillmann, Michael. 2009. Knowledge and presuppositions. *Mind* 118: 241–94.
- Campbell, Joseph K. 1997. A compatibilist theory of alternate possibilities. *Philosophical Studies* 67: 339–44.
- Campbell, Joseph K. 2005. Compatibilist alternatives. *Canadian Journal of Philosophy* 35: 387–406.
- Cohen, Daniel, and Toby Handfield. 2007. Finking Frankfurt. *Philosophical Studies* 135: 363–74.
- Dennett, Daniel. 1984. *Elbow room: The varieties of free will worth having*. Oxford: Oxford University Press.
- Fischer, John M. 1994. *The metaphysics of free will: An essay on control*. Oxford: Blackwell.
- Fischer, John M. 2006. *My way: Essays on moral responsibility*. Oxford: Oxford University Press.
- Frankfurt, Harry G. 1969. Alternate possibilities and moral responsibility. *Journal of Philosophy* 66: 829–39.

¹⁹See also McKenna (2008, 771), who similarly argues, albeit on slightly different grounds, that 'there is more to Frankfurt's argument than the sort of example' on which the literature has tended to focus, and 'inattention to these other aspects of Frankfurt's argument has distorted the force of his case against the freedom at issue'.

- Franklin, Christopher E. 2015. Everyone thinks that an ability to do otherwise is necessary for free will and moral responsibility. *Philosophical Studies* 172: 2091–107.
- Grzankowski, Alex. 2014. ‘Can’ and the consequence argument. *Ratio* 27: 173–89.
- Hawthorne, John. 2001. Freedom in context. *Philosophical Studies* 104: 63–79.
- Jaster, Romy. 2020. Contextualizing free will. *Zeitschrift für Philosophische Forschung* 74: 187–204.
- Kaiserman, Alex. 2016. *Cause and context*. DPhil Dissertation, University of Oxford.
- Kaiserman, Alex. 2017. Necessary connections in context. *Erkenntnis* 82: 45–64.
- Kaiserman, Alex, and Daniel Kocsi. Forthcoming. Review of *Why free will is real*, by Christian List. *Mind*.
- Kratzer, Angelika. 1977. What ‘must’ and ‘can’ must and can mean. *Linguistics and Philosophy* 1: 337–55.
- Kratzer, Angelika. 1981. The notional category of modality. In *Words, worlds and contexts: New approaches in word semantics*, ed. Hans J. Eikmeyer and Hannes Rieser, 38–74. Berlin: de Gruyter.
- Lewis, David K. 1976. The paradoxes of time travel. *American Philosophical Quarterly* 13: 145–52.
- Lewis, David K. 1996. Elusive knowledge. *Australasian Journal of Philosophy* 74: 549–67.
- List, Christian. 2014. Free will, determinism, and the possibility of doing otherwise. *Noûs* 48: 156–78.
- List, Christian. 2019. *Why free will is real*. Cambridge, MA: Harvard University Press.
- McKenna, Michael. 2008. Frankfurt’s argument against alternative possibilities: Looking beyond the examples. *Noûs* 42: 770–93.
- Rieber, Steven. 2006. Free will and contextualism. *Philosophical Studies* 129: 223–52.

Robichaud, Philip, and Jan W. Wieland. 2017. *Responsibility: The epistemic condition*. Oxford: Oxford University Press.

Sartorio, Carolina. 2016. *Causation and free will*. Oxford: Oxford University Press.

Slater, Matthew H. 2005. A contextualist reply to the direct argument. *Philosophical Studies* 125: 115–37.

Szabó, Zoltán Gendler, and Joshua Knobe. 2013. Modals with a taste of the deontic. *Semantics and Pragmatics* 6: 1–42.

Willaschek, Marcus. 2010. Non-relativist contextualism about free will. *European Journal of Philosophy* 18: 567–87.