THE IMPORTANCE OF FRANKFURT-STYLE ARGUMENT

By John Martin Fischer

I reply to the challenges to Frankfurt-style compatibilism about causal determinism and moral responsibility presented in Daniel Speak's paper 'The Impertinence of Frankfurt-Style Argument'. I seek to show how Speak's critiques rest on an 'all-or-nothing' attitude in various ways, and I attempt to defend the importance of Frankfurt-style argumentation in defence of compatibilism.

I. INTRODUCTION

The principle of alternative possibilities (PAP), or something like it, has been at the centre of many of the debates about moral responsibility:

PAP. *S* is morally responsible for something *S* has done only if *S* could have done otherwise.

Certain incompatibilists about causal determinism and moral responsibility invoke (PAP). They contend that causal determinism rules out ability (in the relevant sense) to do otherwise; thus they conclude via (PAP) that causal determinism is inconsistent with moral responsibility.

Whereas some compatibilists have rejected the contention that causal determinism implies 'could not have done otherwise' (in the relevant sense), others have objected to (PAP). Although there are various routes to a rejection of (PAP), one path involves the famous (and in some quarters notorious) 'Frankfurt-style examples'.[1] In these examples, there is a signature structure involving pre-emptive overdetermination; some fail-safe device exists and ensures the actual outcome (i.e., guarantees that the actual outcome will occur) without being triggered or in any way causally influencing the way the actual sequence in fact leads to the outcome in question.[2] In previous work,[3] I have claimed that the Frankfurt-style examples are

---

[1] H. Frankfurt, 'Alternate Possibilities and Moral Responsibility', *Journal of Philosophy*, 66 (1969), pp. 829–39.

[2] For a helpful recent discussion of whether Frankfurt-style examples require a certain sort of pre-emptive overdetermination, see: K. Timpe, 'Trumping Frankfurt: Why the Kane–Widerker Objection is Irrelevant', *Philosophia Christi*, 5 (2003), pp. 485–99.

[3] See, e.g., J.M. Fischer, *The Metaphysics of Free Will: an Essay on Control* (Oxford: Blackwell, 1994); J.M. Fischer and M. Ravizza, *Responsibility and Control: a Theory of Moral Responsibility* (Cambridge UP, 1998); J.M. Fischer, *My Way: Essays on Moral Responsibility* (Oxford UP, 2006).

indeed examples inconsistent with (PAP), and thus that the argument for the in-compatibility of causal determinism and moral responsibility mentioned above can be blocked.

In a fascinating and challenging recent article in this journal, Daniel Speak gives the following Frankfurt-style example:

> Imagine that Paul has insulted John. Paul's insult comes after his own normal and rational deliberation. Nothing suggests the presence of any responsibility-undermining factors at work in the action sequence. However, Frank, a third agent, has all along wanted Paul to insult John. Indeed, Frank has implanted a chip in Paul's brain that allows him to read his intentional states and make changes to them if necessary. Since Paul's intentional states were at each point consistent with his insulting John, Frank has simply continued to monitor Paul's mental content. If, however, Paul had begun to show any inclination not to insult John, then Frank would have activated the chip in such a way as to bring about an inclination in Paul to insult John.[4]

Speak grants (for the sake of the argument) that Frankfurt-style arguments show (PAP) to be false in so far as Paul is morally responsible for his action even though he lacked alternatives with respect to it. But he contends that they are nevertheless entirely irrelevant to the debate between the compatibilists and incompatibilists about the relationship between causal determinism and moral responsibility. That is, Speak contends that even if the examples impugn (PAP), they cannot be used to defend compatibilism about causal determinism and moral responsibility. Speak concludes his article by saying (p. 95) that in the context of debates about the relationship between causal determinism and moral responsibility, 'the famous examples ought simply to be ignored'. Here I shall explain why Speak comes to this conclusion, and I shall seek to show how one could reasonably reject it.

II. SPEAK'S CRITIQUE

As I said above, Speak is willing to concede (for the sake of this discussion) that the Frankfurt-style examples do in fact show the falsity of (PAP), that is, they are cases in which an agent S is morally responsible for an action, although S could not have done otherwise. Of course, this has been a matter of considerable controversy, but Speak does not wish to enter into these debates, and I shall simply follow him here. He does, however, resist the employment of Frankfurt-style examples in seeking to block a certain sort of argument for incompatibilism.

Speak's argument is subtle, but I believe it can be reconstructed roughly as follows. Frankfurt says that the agent S is morally responsible in his examples, despite the fact that he could not have done otherwise, because S would have behaved exactly as he actually behaved, even if he could have done otherwise. That is to say, what explains S's responsibility according to Frankfurt is that S is 'counter-factually stable' with respect to his actual action: he would have done the same

---

[4] D. Speak, 'The Impertinence of Frankfurt-Style Argument', *The Philosophical Quarterly*, 57 (2007), pp. 76–95, at pp. 76–7.

thing, even if he could have done otherwise. But Speak argues that if counterfactual stability (of this sort) is necessary for moral responsibility, then causal determinism is incompatible with moral responsibility. This follows from Speak's contention that causal determinism would rule out the requisite counterfactual stability. Thus according to Speak, the very basis of Frankfurt's explanation of the putative responsibility of the individuals in his examples can be employed as part of an argument for the incompatibility of causal determinism and moral responsibility.

Why would causal determinism threaten counterfactual stability? A part of the argument is this. Arguably, causal determinism rules out 'could have done otherwise'. If so, then the counterfactual supposition that an agent could have done otherwise requires us to move to a possible world where causal determinism is false. But given that causal determinism obtains in the actual world, when we move to a possible world where causal determinism is false, it is hard to know how to evaluate the relevant counterfactual. To be more explicit, suppose an individual S in a causally deterministic world does x. S is counterfactually stable with respect to x only if it is true that if he could have done otherwise, he would still do x. But (on the supposition specified above), the 'closest' possible worlds to the actual world in which the individual could have done otherwise are causally indeterministic. So in order to evaluate the counterfactual, we must ask what the individual who is actually in a causally deterministic world would do in an indeterministic world. And this question is very difficult to answer. For example, what would be the case, if some law which actually obtained never would have been a law? Given the difficulties of evaluation, it simply will not be indisputably true that S would have done x, even if he could have done otherwise. Thus it cannot be established that an agent in a causally deterministic world is counterfactually stable in the relevant way.[5]

The heart of Speak's argument is that counterfactual stability is both required for moral responsibility and incompatible with causal determinism. According to Speak, the Frankfurt-style compatibilist is thus hoist by his own petard: the very element that explains the success of the Frankfurt-style examples is an element in a potent argument for incompatibilism.

III. WHAT FRANKFURT SHOULD HAVE SAID

Frankfurt says that the individual in his examples is morally responsible for what he does because he would have done the same thing, even if he could have done otherwise (counterfactual stability). I believe that Frankfurt should not have said this. In my view, he did not need to offer counterfactual stability as an explanation of the individual's being morally responsible. As Speak notes, Frankfurt offers three different formulations:

(a) ... whatever it was that actually led [him] to do what he did, or that made him do it, would have led him to do it or made him do it even if it had been possible for him to do something else instead.

[5] See also D. Nelkin, 'Irrelevant Alternatives and Frankfurt Counterfactuals', *Philosophical Studies*, 121 (2004), pp. 1–25.

(b) He didn't do it '*because* he couldn't do otherwise'.
(c) The factor that made it impossible for him to do otherwise 'play[s] no role whatever in the explanation of why he did it'.[6]

Speak points out that Frankfurt appears to treat these three formulations as extensionally equivalent. I would contend that Frankfurt should have stuck with (b) and (c) only, and not offered (a) as well. It seems to me that (b) and (c) are (or point to) adequate explanations of the agent's moral responsibility in the Frankfurt-style examples. Perhaps Frankfurt seeks to explain the force of (b) and (c) by offering (a); but (a) is unnecessary, and (as I pointed out above) it gets him into trouble. I would say, employing Bernard Williams' famous phrase, that (a) is 'one thought too many' on Frankfurt's part.[7]

So all Frankfurt really needed to say, and what he should have said, is that the moral responsibility of an individual such as Paul (in Speak's Frankfurt-style example) is not threatened by Frank in so far as Paul does what he does on his own (and without being tampered with by Frank). The point about counterfactual stability is an unnecessary accretion – one thought too many. I do not see why it might be supposed that rejecting (PAP) on the basis of the Frankfurt-examples requires acceptance of the requirement of counterfactual stability; rather, I am suggesting that a proponent of this strategy should say that Paul is morally responsible in so far as he does what he does on his own and not because of tampering by an external factor or agent such as Frank.

Philosophers often assert that an agent in a Frankfurt-style example (say, Paul) is morally responsible because he did what he did 'on his own'. This of course can be unpacked in various different ways. Give Paul all of the features necessary for moral responsibility, including, if you like, indeterminism (of a certain sort) in the deliberative process. Paul is, in fact, just like an agent all parties would agree is morally responsible, save for one difference: Paul is in a world containing another agent (or process) stipulated to be 'causally irrelevant' to any of Paul's deliberations or actions. Such a non-factor could not make a difference to Paul's being morally responsible, could it? Metaphysically, what explains Paul's moral responsibility is all of the necessary conditions which were put in earlier. Dialectically, what allows one to continue saying that Paul is morally responsible is that the would-be intervener does not do anything.

It seems to me that counterfactual stability plays no explanatory role – it is not even mentioned or (as far as I can see) implicitly invoked; rather, reliance is placed on the idea of Paul's doing what he does 'on his own' (and not because of Frank's tampering). Given the availability of this sort of Frankfurt-style strategy, I believe it is clear that a Frankfurt-inspired rejection of (PAP) does not require counterfactual stability; I shall further argue for this position in §V below when I address Speak's response to this sort of suggestion.

Of course, from the fact that it was unnecessary for Frankfurt to invoke a requirement of counterfactual stability for moral responsibility in the context above it does not follow that such a requirement does not exist. I do not, however, believe that counterfactual stability is a genuine requirement for moral responsibility. Interpreted as Speak interprets it, this is manifestly too stringent a requirement, since Frankfurt-type cases can be constructed in which the agent appears responsible but it is not present. Suppose some form of indeterminism in deliberation in a Frankfurt-style case.[8] Given the indeterminism, there is no counterfactual stability; such stability entails that were we to replay the relevant stretch of history, the agent would always decide as he did, and indeterminism rules this out. A requirement of counterfactual stability seems to expunge moral responsibility even apart from the special context of Frankfurt-style examples. Thus not only would I contend that a proponent of Frankfurt-style compatibilism need not invoke counterfactual stability, I would also point out that this requirement is independently implausible.

IV. SPEAK'S REPLIES

To his credit, Speak anticipates the view I have suggested above (according to which one does not rely on counterfactual stability), and he offers various replies. It will be useful to consider them carefully. First, he points out (p. 92) that (a) may well be necessary for (b) and (c); he wonders how one would establish (b) or (c) without invoking (a). He does however appear to acknowledge (pp. 92–3) that one could adopt some sort of non-reductionist, 'causal impact' (or, presumably, 'production') view of causation on which (a) could in principle be separated from (b) and (c).

His reply (pp. 92–3) is as follows:

Frank 'plays no role' in Paul's insulting behaviour (or Paul does not do the insulting *because* of Frank) if no causal force is transmitted between them. Any appeal to this sort of argument, however, will not extricate the Frankfurtian compatibilist from the impertinence I have identified. This is because the truth of causal determinism would entail that the past (together with the laws) *does* transmit causal force to any future choice or action.

... There is no plausible way of extending the implications of Frankfurt-style cases into the context of determinism so as to provide any real resistance to the standard incompatibilist argument. With respect to the compatibility debate, then, the famous cases are an irrelevant distraction.

Speak also considers (pp. 93–4) a related objection that has it that the Frankfurt-style examples provide a useful first step in the defence of compatibilism, even if they are not themselves decisive components of such a defence:

The eirenic reply to this point would be to accept it, admit that my initial claims are somewhat rhetorically inflated, but emphasize the constraints that my argument has placed on a completely satisfying Frankfurtian compatibilism. I am tempted by this

[8] See, as just one example, A. Mele and D. Robb, 'Rescuing Frankfurt-style Cases', *Philosophical Review*, 107 (1998), pp. 97–112.

reply. But I shall resist the temptation. This is because I think it would unjustifiably concede too much to the compatibilist. The guiding insight of my approach has been that (PAP) must be false *in the right way* for the compatibilist to get leverage. More specifically, (PAP) must be false in a way that can be exploited under determinism. The Frankfurtian way is not, I have argued, of the right sort in this respect.

V. THE IMPORTANCE OF FRANKFURT-STYLE ARGUMENT

I believe that the response to Speak's two replies here stems from a certain way of understanding the dialectic strategy of Frankfurt-style compatibilism.[9] The key is not to think of the strategy as proceeding in a precipitous way; it is not a one-step argument, but involves at least two important steps or stages. The first step is that the examples show that (PAP) is false, so moral responsibility does not require alternative possibilities. Thus causal determinism cannot be thought to rule out moral responsibility *simply* in virtue of eliminating access to alternative possibilities. Of course, if causal determinism obtains, then alternative possibilities are ruled out *in a certain way* (i.e., via a causally deterministic actual sequence). The second step of the argument is this: if causal determinism does not rule out responsibility by eliminating access to alternative possibilities, why exactly would it be thought to rule out moral responsibility by *eliminating alternative possibilities in this specific way* (i.e., via a causally deterministic actual sequence)? That is, if it is no problem for moral responsibility that the agent does not have access to alternative possibilities, why would it be a problem that the agent lacks access to alternative possibilities because of a causally deterministic actual sequence?

Of course, there might be something problematic and challenging for moral responsibility in a causally deterministic actual sequence. In other words, it might be that causal determination in the actual sequence rules out moral responsibility for some reason other than that it eliminates access to alternative possibilities. Perhaps actual-sequence determination is incompatible with 'initiation' or 'sourcehood' or 'creativity' or ....[10] This claim is important, and raises difficult issues that must be addressed. But here it is important to note that the debate has shifted to a different dialectical terrain – we are here not asking about the threat posed by causal determinism to alternative possibilities; rather, we are asking whether causal determinism *directly* rules out moral responsibility (that is, whether causal determinism threatens moral responsibility in itself and apart from eliminating access to alternative possibilities). Although in the end I would argue that direct incompatibilism fails, I think it is a perfectly plausible position whose appeal I can understand. What is puzzling to me is how one could accept the first step of the Frankfurt-style compatibilist's argument (i.e., that causal determinism does not threaten moral responsibility by eliminating access to alternative possibilities) but claim that it does threaten moral responsibility by involving a certain path to the ruling out of

[9] I seek to explain this sort of strategy in 'Frankfurt-Style Compatibilism' and 'Free Will and Moral Responsibility', both of which are reprinted in *My Way*.

[10] See my *The Metaphysics of Free Will*, pp. 147–54; Fischer and Ravizza, *Responsibility and Control*, pp. 151–69; and my 'Frankfurt-Style Compatibilism'.

alternative possibilities – by being a *certain way of eliminating alternative possibilities*. If the elimination of alternative possibilities *in itself* is not problematic, and if the path *in itself* (apart from leading to the elimination of alternative possibilities) is not the source of the problem, how can the problem stem from the fact that causal determinism involves a certain way of eliminating alternative possibilities?

If they are interpreted in this more complex way, it can be seen that the Frankfurt-style examples help a certain sort of compatibilist to make considerable progress. The argumentative strategy of the compatibilist involves more than one step, and it is admittedly not a knockdown argument. But why suppose arguments here must be (maximally) simple and knockdown? My point applies to Speak's replies in this way: Speak insists that the Frankfurt examples are a different sort of context from that of causal determinism, and (relatedly) that (PAP) must be shown to be false 'in the right way' to be any use to the compatibilist. One can think of Speak's view in the following way. In the Frankfurt-style examples, some factor that operates in the alternative sequence, but not in the actual sequence, renders it true that the agent lacks access to alternative possibilities. In contrast, in a case where it is explicitly assumed that causal determinism obtains, some factor (or factors) operate in the actual sequence, and (we can assume) render it true that the agent lacks access to alternative possibilities. So Frankfurt-style examples contain a freedom-undermining factor only in the alternative sequence, whereas causal determinism involves actual-sequence factors that undermine freedom. One cannot therefore straightforwardly apply lessons putatively learned from Frankfurt-style examples to cases in which it is assumed that causal determinism obtains.

I grant this point; one cannot straightforwardly and in one step go from the Frankfurt-style examples to a conclusion about cases in which causal determinism is assumed to obtain. But it does *not* follow that one cannot employ the more nuanced argumentative strategy I sketched above. Here the lesson of the Frankfurt-examples is not straightforwardly applied to contexts of causal determinism, but the argument proceeds in (at least) two steps. Further, it is not alleged by the Frankfurt-style compatibilist that the strategy is *knockdown* or *decisive*. Given a proper understanding of the structure and nature of Frankfurt-style compatibilism, Speak's conclusions that 'the famous cases are an irrelevant distraction', that 'they make no contribution to the compatibilist issue', and that they 'ought simply to be ignored' can be seen to be extreme and certainly unwarranted.

Finally, I shall consider an objection to the argumentative strategy I have suggested on behalf of the compatibilist. I have accepted (for the sake of argument) that an agent such as Paul in a Frankfurt-style case is morally responsible because of (b) – what makes Paul morally responsible is that he did not behave as he did because he could not have done otherwise. Arguably, in a causally deterministic context Paul does indeed behave as he actually behaves because he could not have done otherwise. Thus it appears that the one factor that saved Paul's moral responsibility in the first place (in spite of there being no alternative possibilities open to him) is *missing* in a deterministic scenario. Why, then, should we think that Paul is morally responsible in a causally deterministic scenario?[11]

[11] I owe this point to David Robb.

The dialectical issues here are admittedly subtle. My reply is that it is *not* the case that the one factor that saved Paul's moral responsibility was that he did not do what he did because he could not have done otherwise. Rather, the idea above was this: we can invoke Frankfurt-style examples to show that in principle it is possible for an agent to be morally responsible, even though he lacks alternative possibilities; more specifically, we invoke such examples at a point in the dialectic which does not presuppose causal determinism, and thus the conclusion (that an agent could be morally responsible even in the absence of access to alternative possibilities) should be acceptable even to someone inclined to (or committed to) incompatibilism about causal determinism and moral responsibility. Here it is not being claimed explicitly or implicitly that Paul's not doing what he does because he could not have done otherwise is a factor that 'saves his responsibility', or in the absence of which he would not be morally responsible. Rather, the idea was that we can invoke (b) at a particular point in the argument in a way that is dialectically productive – it helps us to make progress in a non-question-begging way. More specifically, the idea is that even someone inclined towards incompatibilism should agree that if causal determinism is false and the relevant agent did not have access to alternative possibilities (because of the existence of a Frankfurt-style counterfactual intervener), the agent might nevertheless be morally responsible – given that the actual sequence proceeds in the right way. So the strategy I suggest does not suppose that (b) specifies a factor that would makes a metaphysical difference (in the presence of which an agent is morally responsible and in the absence of which he is not); rather, (b) specifies a factor that is dialectically fruitful in the way sketched above. Of course, a lot then depends on the *second* stage of my argumentative strategy.

VI. CONCLUSION

It seems to me that Speak's critique of Frankfurt-style compatibilism suffers from a certain kind of 'all-or-nothing' attitude. He appears to assume that the proponent of Frankfurt-style compatibilism simply asks us to consider the examples, and precipitously to conclude that causal determinism is clearly and indisputably compatible with moral responsibility. But I have shown that the Frankfurt-style compatibilist need not pursue this sort of strategy, a strategy which should in any case be highly suspicious to anyone who has thought carefully about free will and moral responsibility (or about any other difficult and contentious philosophical issue). The Frankfurt-style compatibilist can offer an argument *in stages*, and one that does not purport to be impossible to reject (especially by anyone antecedently committed to or strongly inclined to incompatibilism).[12]

University of California, Riverside