



# Alternative Possibilities and Causal Overdetermination

**Ferenc Huoranszki**  
Central European University

BIBLID [0873-626X (2017) 45; pp. 193–217]

DOI: 10.1515/disp-2017-0004

## Abstract

This paper argues against dismissing the Principle of Alternative Possibilities merely on the ground of so-called Frankfurt-style cases. Its main claims are that the interpretation of such cases depends on which substantive theory of responsibility one endorses and that Frankfurt-style cases all involve some form of causal overdetermination which can be interpreted either as being compatible with the potentially manipulated agent's ability to act otherwise or as a responsibility undermining constraint. The paper also argues that the possibility of such scenarios can support the truth of classical compatibilism as much as the truth of semicompatibilism.

## Keywords

Freedom, possibilities, abilities, overdetermination, preemption.

The principle of alternative possibilities as I here understand it consists of two theses. The first thesis is that an agent can perform an act freely only in such conditions in which she is able to avoid doing what she does; either by doing something else or by simply refraining from acting. The second is that an agent cannot be responsible for her behavior unless she acts freely and that she cannot be responsible for the future effects of her behavior unless they are the foreseeable potential consequences of something that she did freely. This means that responsibility understood as a condition of the normative evaluability of agents' behavior always requires the ability to avoid doing what one does; or to do what one has actually failed to do.

The truth of the principle of alternative possibilities (PAP for short) has always been a matter of controversy. Some believe, for instance, that the ability to do otherwise is incompatible with the

truth of physical determinism. But while the truth of physical determinism is an open question, it cannot be an open question that we are often responsible for what we do: responsibility is a sort of Moorean fact. Consequently, alternative possibilities cannot be relevant to agents' freedom and responsibility. Others hold—like the so called 'necessitarians' about free will did—that having an ability to do otherwise could only weaken the agent's capacity to exercise rational self-control and hence it cannot be relevant to her freedom and responsibility.

Surprisingly, contemporary debates over PAP seem to be rarely motivated by such general considerations about the nature of agents' freedom. Rather, they concern some peculiar cases which can allegedly be sufficient to prove that alternative possibilities are irrelevant to the issue of freedom and responsibility. The original example was introduced in a highly influential paper by Harry Frankfurt in the late sixties.<sup>1</sup> Subsequently other philosophers have proposed different variants of such 'Frankfurt-style cases' (FSC for short) with the ambition to prove that an agent's ability to avoid doing what she does is *not* a condition of responsibility—neither, in fact, of freedom—irrespective of which substantive theory of freedom and responsibility one endorses.

In this paper I am going to argue that this ambition cannot be satisfied. Whether or not the proposed examples can prove the irrelevance of alternative possibilities still depends on some prior substantive theory about the nature of freedom and the conditions of responsibility in general. Any theory of freedom and responsibility that requires PAP can respond to FSCs either by asserting that, in some practically relevant sense, an agent in FSC is able to choose and act otherwise; or by denying that she can be responsible for her behavior even if there is no sense in which she can do otherwise. Consequently, FSCs cannot provide us with any *further* reason to reject PAP independent of our substantive account of freedom and responsibility, which may or may not entail it.

<sup>1</sup> Reprinted in Frankfurt 1988.

## 1 The structure of Frankfurt-style cases

Though Frankfurt was certainly not the first philosopher who argued against PAP as capturing a necessary condition of moral responsibility, contemporary debates over the validity of the principle are dominated by discussions of cases inspired by his argument. Earlier contenders of PAP grounded their claim about the irrelevance of alternative possibilities in some theory of moral (or metaphysical) freedom, which does not require—or might even be in conflict with—PAP. In contrast, Frankfurt in the paper that sets out to reject PAP does not offer any account of agents' freedom as a condition of their responsibility. He does this only later. His original argument against PAP does not rely on it.

This apparently accidental historical fact has some consequences in substance. Frankfurt claims that no matter how the ability to do otherwise is interpreted his argument against PAP should stand. But if the relevant sense of abilities can depend on some substantive account of the conditions of an agent's responsibility, then one can argue that FSCs can show at most the irrelevance of certain kinds of alternatives while not affecting the truth of PAP in general. And as we shall see, this seems precisely to be the case.

Frankfurt's argument against PAP is based on examples. In order to prove the irrelevance of alternative possibilities such examples must satisfy two conditions. First, we must have the intuition that the agent in the example is responsible for her behavior. Second, it must be clear that the agent is unable to act otherwise so that she lacks alternative possibilities in the sense which might initially—i.e. before the argument's conclusion is granted—seem to be relevant to responsibility. Since intuitions both about agents' responsibility and about the relevant notion of ability are unstable, the exact formulation of FSC is a contentious matter.

Instead of rehearsing Frankfurt's well-known story or introducing my own, let me just summarize the general structure of the examples. They look like this:

- (1) An agent M (the manipulator) has the power to neurologically manipulate another agent S's choices.

- (2) M wants S to perform a particular kind of action A in circumstances C.
- (3) M lets S to make his own choice unless he recognizes that S is about to choose to avoid doing A.
- (4) If S is about not to choose to do A, M intervenes and causes S to choose to do A.
- (5) If this happens, S is not responsible.
- (6) If, however, S chooses to do A 'of his own', he is responsible.
- (7) Conclusion: PAP is not a condition of responsibility because although, given (4) S is unable to avoid doing A, he is still responsible if he chooses to do A 'of his own'.

In the remaining of this section, I shall discuss some concerns about the possibility of the scenario as described by Frankfurt and explain how such concerns led to some modifications of the original example.

An obvious difficulty about the possibility of FSCs concerns the concept of choice required by Frankfurt's example. After all, in the case when the manipulator remains inactive the ground of S's responsibility seems to be that he 'makes his own choice' as opposed to his choice being caused by M's intervention. Does this mean that S's 'own choice' is not caused by anything? Or, perhaps, that it is not caused by any prior event, but it is caused by him as an agent? The latter would presuppose an agent-causal account of freedom while the former requires that choices be uncaused events. However, in either case, how is it possible that *the very same type of event*—that is the agent's choosing this or that—can also occur as a result of some external causal intervention? If one thinks that choices cannot be caused, or at least they cannot be caused by any antecedent processes or events, then FSCs simply cannot get off the ground.

Hence, in order to make sense of a FSC, we must assume either that an agent can make a *directly causally manipulated choice* (indirect manipulation, like manipulating an agent's choice by giving reasons or offering new options is obviously irrelevant in the present context); or we must introduce the possibility that some sort of events

in an agent's mental life, even if they are not his genuine choices, can function similarly to those choices. For instance, we can say that the formation of intentions to act can be governed by choices\*; that what an agent chooses\* is not unintelligible given her reasons; and that what she does depends (actually or counterfactually) on what choices\* she makes.

Then, even if an agent's *own* choice cannot be caused, we can interpret FSCs as describing a situation in which the agent might be caused to make a choice\*; where choice\* is a mental occurrence with a role sufficiently similar to the one that 'normal' choices play in an agent's life. And we should then reformulate (4) accordingly: if S was not to make *her own uncaused or agent's caused choice* to A, then an alternative (manipulative) mechanism would *cause her to choose\** to A. That all this makes clear sense might be controversial, of course, but we need to live with it if we want to reconcile FSCs with the possibility that genuine choices are not—either deterministically or indeterministically—(event-)caused.<sup>2</sup>

Another difficulty concerns the identification of the piece of behavior for which an agent can or must be responsible. Some argue that even if the potentially manipulated agent is indeed unable to avoid doing *the kind of* action she does, she can be responsible *for a particular act* only if she can avoid it. Alternatively, one can argue that a merely *potential* manipulation does not undermine the agent's ability to act otherwise; and it is for this reason that we hold S in a FSC responsible.

In fact, most replies to FSC are some combination of these two kinds of response. They argue that Frankfurt's original example underspecifies the action for which the agent is responsible as well as the conditions in which the potentially manipulated agent acts; and that the proper description of the cases faces an unanswerable dilemma: *either* it seems false that the potentially manipulated agent cannot do otherwise (in the sense relevant to her freedom and responsibility); *or* that she is not responsible for the sort of action she does.<sup>3</sup>

<sup>2</sup> Particularly because we have good reasons to believe that choices are not caused even if physical determinism is true. The truth of physical determinism does not entail universal causation. See Dennett 2003.

<sup>3</sup> Variations of this argument are found in van Inwagen 1978 and Ginet 1996.

A related, but more serious difficulty about the possibility of genuine FSCs concerns the timing of the manipulator's decision to intervene rather than the identity of actions for which S may be responsible. Suppose  $t_4$  is the time by which M wants S to do A and  $t_3$  is the time by which S must make his choice to do A at  $t_4$ . Then there must be an earlier time  $t_2$  which is the last time for M to interfere and cause S to choose at  $t_3$  to do A at  $t_4$ . But paying more attention to the temporal order of events in FSCs betrays the following problem with Frankfurt's original example.

It seems that if the argument does not want to presuppose the truth of its conclusion—stating that agents can be responsible for actions which they cannot avoid—then it must allow that whenever we hold S responsible we (partly) do so on the ground that at any time *before* he makes his choice, he still has the ability to choose (and do) otherwise. What typically happens at  $t_3$  is, after all, that the agent arrives at some deliberative conclusion, and if the outcome of a deliberative process is unavoidable already during the process of deliberation, then S is unable at  $t_1$  (that is at any time before she makes her choice) to choose otherwise at  $t_3$  anyway—irrespective of the presence or absence of M.

But then we face the following dilemma. On the one hand, if until the time of decision (or coming to a deliberative conclusion) it is *not* open to S both to choose to do A and to choose to avoid doing A later, then FSC cannot *prove* the irrelevance of the ability to choose otherwise; it must simply assume it. On the other hand, if until  $t_3$ , that is until the time by which the deliberative conclusion is reached, it is open to S to arrive at another conclusion and thus choose otherwise, the kind of intervention required by FSC is simply impossible. For there is no time when the potential intervener can be assured of the necessity of intervention. Hence, whenever free and responsible, S must be able to choose otherwise all along.<sup>4</sup>

This problem besets Frankfurt's original example, which requires that the intervention into the agent's choice be a response to what the agent does or is supposed (by the intervener) to do before he makes his choice. In that example, the manipulative intervention must be

<sup>4</sup> Versions of such considerations are offered by libertarians like Widerker (1995), Ginet (1996), Kane (1996).

triggered as a response to something that S (non-intentionally) does, and which must somehow indicate to the potential manipulator that S is not going to choose to do A. However, as we have just seen, one cannot simply assume that there is a moment during the deliberative process after which a free agent cannot change his mind about what to do. Hence there is no way for M to respond to S's behavior before the latter reaches a deliberative conclusion and hence chooses to do A. More precisely, if we assume that S can be responsible even if he is unable to choose otherwise at  $t_2$  whenever some of his pre-choice states at  $t_1$  'indicates' that he *will* choose to do A at  $t_3$ , then the example cannot prove, without circularity, the irrelevance of alternative possibilities.

But the original FSC is perhaps only a rough approximation. Perhaps we can provide some more complex account of S's deliberative process, which is compatible with the assumption that there is some such state of the agent at  $t_2$  that precedes S' choice at  $t_3$  and indicates her future choice without causing it. The occurrence of that earlier state might be necessary for S' choosing at  $t_3$  to avoid doing A, and hence incompatible with her not choosing to do A, without causing his choice. Such states—or their absence—can then provide M with the necessary 'prior sign' to intervene before  $t_3$ . In this case although S has an ability to issue or to avoid issuing the relevant sign, it seems implausible that we should hold him responsible on that ground. For those alternatives are not 'robust enough' to ground the agent's freedom and responsibility.<sup>5</sup>

However, such attempts to revise the original example look unpromising for at least two reasons. First, in order to make such cases possible, we must introduce several dubious assumptions about the nature of deliberative process and its relation to an agent's choices. For instance, we need to assume that some of S's neural states are such that they can somehow 'block' the possibility that S chooses to avoid doing A after  $t_2$  even if he *can* continue to deliberate about that option. It is hard to imagine how this is possible. Moreover, this is exactly the kind of assumption that many who grant the relevance of

<sup>5</sup> For an attempt to construct FSCs on these lines see for instance Pereboom 2001. About the concept of robustness see more recently Fischer 2012: 39; as well as many of his earlier papers like Fischer 2002. For criticism see e.g. Nelkin 2011.

alternative possibilities would reject.

Second, and more importantly, it is unclear why the alternatives allowed by these examples are not 'robust enough' for the agent's responsibility. The potentially manipulated agent can, after all, deliberate. It is certainly true that he is not considering whether or not to issue the relevant sign; if not for other reasons, just because he is unaware of that possibility. But an agent can be unaware of the fact that certain conditions must obtain in order to make his own choice, nonetheless those conditions might be among the robust conditions of his responsibility.

An agent might be unaware of the fact that, given the supervenience of her deliberations and choices on her neural processes, she cannot make a different choice without being in a different sort of neural state; and even if she is aware of this *in abstracto*, she is certainly not aware of which states they are, so she cannot directly choose to be in that state. Nonetheless, it should be common ground among every participant in this debate that if going through a somewhat different neural process is a necessary condition of making a different choice, then the ability to go through that neural process must be a robust enough condition of freedom, even if the agent does not, and cannot, make any choice about it. Hence what matters is not whether the agent can make a choice about issuing 'prior signs', but that the *common cause* of issuing it and making the corresponding decision has or has not been avoidable.

However, there is a more interesting, and *prima facie* more promising way to modify Frankfurt's original example. One can simply reply that the whole issue about the timing of manipulative intervention is irrelevant. The original example as well as many of its proposed modifications presuppose that the manipulative intervention must be triggered as a response to the potentially manipulated agent's behavior. But no such triggering is needed if the manipulative intervention is already on its way independent of how the agent deliberates or of which kind of causal process (if any) leads to his choice (or choice\*). The intervention might occur, if it does, at the same time at which the agent reaches a deliberative conclusion and hence makes her choice. This means that we have to reinterpret premise (4) of the original argument. But we can appreciate the significance of this proposal adequately only if first we investigate in more detail the



causal structure of the situations involved in FSCs.

## 2 Causal overdetermination

As Frankfurt himself makes it clear, his example against PAP is based on the observation that an agent's behavior—like any other event—might be *overdetermined*. However, an event (as well as a choice or an action) can be overdetermined in several different senses, and not every kind of overdetermination is compatible with the agent's responsibility. I suggest then that the most promising way to understand the debates over the possibility of genuine FSC is to investigate the nature and consequences of overdetermination which must be involved in such cases.

There are several distinct forms of qualitative overdetermination, but here it is important for us to distinguish only two basic variations.<sup>6</sup> An effect might be *actually* causally overdetermined; or it might be only *potentially* overdetermined. Roughly, again, *actual* overdetermination occurs when two (or many) independent causes play an actual role in the production of their common effect, where any of those causes would in itself be sufficient to produce the effect. *Potential* overdetermination occurs in those circumstances in which, although only one independent cause produces the effect, had that actual cause failed to bring it about, another alternative cause would have produced it.

Philosophers disagree over the issue whether or not effects can ever be actually qualitatively overdetermined. For although it is hard to deny that there are cases in which the *same sort of* effect can be brought about by independent causes, it is a contentious matter whether we should allow that *the same particular event* would occur if one of the causes were absent (in the actual overdetermination case); or if another cause were operating than what has actually been operating (in the potential overdetermination case). And, as we have seen already in the previous section, such issues concerning the identity of effect-events can affect the interpretation of FSC.

Philosophers also disagree terminologically (as they often do).

<sup>6</sup> For a useful summary of the distinct kinds of overdetermination see Mackie 1974 and Lewis 1986.

On some accounts, potential overdetermination is not a form of *causal* overdetermination at all, since an effect can be overdetermined in the potential sense only if it has one single actual cause. Some claim that preemptive overdetermination is not a form of causal overdetermination.<sup>7</sup> But this disagreement about terminology has no consequence in substance. The important point is that an intuitively convincing FSC must describe cases of *potential or pre-emptive overdetermination*.<sup>8</sup>

What makes us think that the agent in FSC can be responsible—and perhaps also free as Frankfurt later suggests—is that the potential cause, that is the potential intervention or manipulation, *plays no actual (causal) role* in the *production* of the agent's choice and behavior; neither does it explain why the agent acts in the way she does. In those cases in which the actual presence or absence of some factors do play some causal role in the production of the agent's behavior it is *never obvious* that we should hold the agent responsible; unless we grant already, and independent of any FSC, that agents can be responsible even if their behavior is subject to actual manipulation.

My claim about the importance of the mere potentiality of overdetermination can be challenged on the ground that, according to Frankfurt, in his example the 'circumstances *actually played a role in bringing it about* that he did [something which was impossible to avoid], so that it is correct to say that *he did it because* he could not have done otherwise' (Frankfurt 1988: 11, my emphasis).<sup>9</sup>

However—as I have argued earlier<sup>10</sup>—it is difficult to reconcile this with Frankfurt's own description of the case as an arrangement which 'ensures that a certain effect will be brought about by one or the other of the two causal factors, *but not by both together*. Thus the backup factor *may contribute nothing whatever to bringing about the*

<sup>7</sup> See Hitchcock 2011. According to Hitchcock's terminology, potential overdetermination is causal preemption, not causal overdetermination. I follow here the tradition that originates in Mackie 1974, which does consider cases of potential overdetermination as a form of causal overdetermination. As should be obvious, however, not much depends on the choice of terminology.

<sup>8</sup> As has been claimed both by Frankfurt and his critics like e.g. Kane (2003).

<sup>9</sup> As it has been emphasized, for instance, by Funkhouser (2009).

<sup>10</sup> In Huoranszki 2011: 83.

*effect* whose occurrence it guarantees' (Frankfurt 1988: 96; my emphasis).<sup>11</sup> Most subsequent interpretations of FSCs, in fact, agree that cases of actual overdetermination cannot constitute successful FSCs.<sup>12</sup>

But there are more important reasons to reject the idea that actual overdetermination of the agent's choice is compatible with his freedom, which are neither exegetical (noting that Frankfurt seems to be committed to that view) nor merely statistical (mentioning that most of his followers have understood the case in this way).

First of all, most would agree that S cannot be free and hence responsible if his choice\* has been brought about *exclusively* by M's manipulative intervention (so that he could not even make 'his own choice'). Now if actual overdetermination were compatible with responsibility, then we should assume that the agent, by 'making his own choice' and thereby simply adding some *extra* causal factor, could make himself responsible for something that has already been brought about anyway. But this is more reminiscent of a sort of wishful thinking than the kind of control that we usually assume to be necessary for responsibility and freedom.

Second, and most importantly, any FSC that involves actual overdetermination *must begin with* the assumption that an agent's freedom and responsibility does not require the ability to choose and do otherwise. And this seems to contradict the general argumentative strategy used in FSCs. For the intuition that drives FSCs is the following: if in those circumstances in which the alternative causal process is absent the agent is responsible because he 'makes his own choice', then he must also be responsible in circumstances in which certain factors 'may contribute nothing whatever to bringing about the effect whose occurrence it guarantees' (Frankfurt 1988: 96, my

<sup>11</sup> And further 'But there may be circumstances that make it impossible for a person to avoid performing some action *without those circumstances in any way bringing it about that he performs that action*. It would surely be no good for the person to refer to circumstances of this sort in an effort to absolve himself of moral responsibility for performing the action in question. For those circumstances, by hypothesis, *actually had nothing to do with his having done what he did*' (Frankfurt 1988: 9, my emphasis.)

<sup>12</sup> Fischer and Ravizza, for instance, say that the FSC requires preemption because we can hold agents responsible only if the backup conditions are not causally efficacious in bringing about the action (Fischer and Ravizza 1998: 159).

emphasis). But if we begin with the assumption that an agent can be responsible even when her choice and action *has actually been brought about* by some intervention, then the issue about the relevance of alternative possibilities does not even arise.

Thus FSCs with actual overdetermination can do no more than *illustrate* some cases of responsibility without PAP given that *we are already committed* to some positive accounts of responsibility which does not require the agent's ability to do otherwise. FSCs cannot be used as an independent argument against PAP. For any theory of freedom that allows its compatibility with actual overdetermination of choices the role of FSCs seems to be redundant. Consequently, in order for FSCs to serve as a theoretical motivation of the rejection of PAP, they must involve some form of potential causal overdetermination.

Most potential overdetermination cases are cases of *early* preemption. This means that the completion of the preempted, potential causal process CP2 is 'cut off' by the preempting process CP1 sometime before the occurrence of the effect. In the original FSC, something the agent does (non-intentionally, of course) 'cuts off', silences, or prevents the potential intervention. Since preemption is early only if the causal process which actually produces the effect can also cut off, and hence abort the completion of, the alternative process sometime before the effect is produced, early preemption presupposes that the preemptive and the preempted processes can *interact before* the (potentially, but not actually, overdetermined) effect occurs.

However, FSCs might involve *late* rather than *early* preemption.<sup>13</sup> Late preemption does not require any such interaction between the preempting and preempted causal process. It requires only the operation of two independent causal processes, which 'compete', as it were, for causing the same effect; but which do not jointly cause it. Frankfurt's own example is a case of early preemption, since in the original example M's potential intervention is actually 'cut off' by something which S does before his choice and which can 'ensure' M that S is going to choose, 'of his own', to do A later. Hence the process of S's deliberation which leads to his choice also blocks, in a sense, M's potential intervention. However, it can be argued that

<sup>13</sup> On the distinction between late and early preemption see Lewis 1986: 200–212.

potentially triggering—or actually silencing—M's intervention is not an essential part of FSC. For the intervention can be conceived of the model of late, rather than early, preemption.

Suppose  $CP_A$  is an agent's *indeterministic* deliberative process which might result in the agent's choosing at  $t_2$  to do A at  $t_3$ . Suppose  $CP_M$  is an alternative *deterministic* causal process, uncontrolled or unaffected by the agent himself, which can result in causing him to choose at  $t_3$  to do A. Suppose, finally, that if the agent's indeterminate deliberative process actually concludes with her choosing to do A, then his own choice can 'trump' the alternative causal process in the sense that it does not cause the choice. It might seem then that the agent can 'make his own (non-determined) choice' at  $t_3$  to do A at  $t_4$ —and hence he is free and responsible—even if he is unable to choose to avoid doing A. For if he had not made his own (non-determined) choice, then it is the alternative causal process  $CP_M$  which would have caused him to make that choice exactly at the same time.

This is the basic structure of the version of FSC proposed in an influential paper by Alfred Mele and David Robb (Mele and Robb 1998). This version of FSC looks perhaps the most promising argument against the relevance of alternative possibilities. But the importance of the trumping scenario, I shall argue, is not that it proves the irrelevance of alternative possibilities, but rather that it helps diagnose the main problem with FSCs in general.

### 3 Trumping preemption and the indeterminacy of deliberation

I said that the overdetermination in the Mele-Robb case involves *a sort of* trumping, though the concept of 'trumping preemption' was introduced by Jonathan Shaffer only somewhat later as an objection to counterfactual analyses of causation.<sup>14</sup> Shaffer calls 'trumping' causal scenarios with the following features. Suppose there are two independent, non-interacting causal processes CP1 and CP2, both of which have the power to directly produce the occurrence of an event E exactly at time  $t$ . However, suppose further, that one 'cause's

<sup>14</sup> See Shaffer 2000. Of course, similar sort of examples appeared before Shaffer named them, particularly in Ehring 1997.

power' can override the other's power in the following sense: in any situation, in which *both* CP1 and CP2 were launched so that each were ready to produce E's occurring at *t* in the absence of the other, *only* CP1 would cause E. It is in this sense that CP1 can 'trump' CP2. Importantly, CP1 does not cause E by aborting CP2. It preempts *only CP2's causing E* just by causing the effect itself.

Shaffer argues that trumping is physically possible and that it is a form of preemption. Both claims are controversial.<sup>15</sup> But we can put such controversies aside here and assume that trumping preemption is conceptually possible. The interesting question is whether or not FSC can be interpreted on the model of this kind of preemptive overdetermination; and how this interpretation can prove the irrelevance of alternative possibilities for an agent's freedom and responsibility. For it seems obvious that *some* form of trumping is necessary for the Mele-Robb example. The potentially manipulated agent's deliberative process must have the power to 'override' the power of the alternative process in the sense that the former can preempt the latter's completion just when it ends by directly causing (or at least resulting in) the agent's choice to do A.

Despite this important structural similarity, however, this version of FSCs differs from standard trumping preemption cases in a crucial respect. In those cases as described by Shaffer, CP1's causing event E trumps CP2's causing E. But the standard cases are deterministic, while in the case described by Mele and Robb CP1 is nondeterministic in the sense *that it might cause E\* at t<sub>2</sub> instead of causing E*. As just said, trumping cases are founded on the idea that one cause's power to produce E at *t* somehow overrides, and hence can trump, another cause's power to produce E at *t*. But in the Mele-Robb case, although the agent's (or his deliberative process's) power to choose action A at *t<sub>3</sub>* overrides CP2's power to produce the same

<sup>15</sup> The standard example for trumping is an order simultaneously issued by a sergeant and a major. The soldiers are supposed to be caused to act upon the major's order, even if it is the same as the sergeant's order. However, it is unobvious why not *both* orders are obeyed; and anyway, what sort of causal structure obeying orders has. Shaffer also provides physical examples, the actual occurrence of which would, however, contradict the general physical principle that exercising forces must always actually influence the momentum of bodies (unless we add them together as counteracting forces, but that is irrelevant for the example).

choice, CP2's does also have the power *to prevent* the agent's (or his deliberative process) from causing his choosing to avoid A at  $t_3$ . Consequently, in the Mele-Robb case, unlike in the original trumping preemption cases, CP2 does not have the role of bringing about E in case CP1 would not; rather its role is to *prevent CP1's power* from bringing about E\*. And this makes the example problematic. For actively preventing the agent from making a certain choice obviously undermines his freedom and responsibility.

Moreover, in order for the example to work, we need to imagine a possible scenario in which a causal process CP2 is '*weaker*' than CP1 in the sense that whenever CP1 is launched and completed CP2 *cannot* cause E; but at the same time, CP2 is '*stronger*' than CP1 in the sense that if CP1 were to cause E\* rather than E, then CP2 *could* prevent it just at that point by causing E instead. Though Mele and Robb claim that they can describe analogous cases outside the context of choice, in fact, it is very hard to imagine how any real life physical system can be configured with such a complex trumping structure. For as we have just seen, Mele and Robb's case involves a sort of *double trumping*. CP1 can trump CP2 if doing A is chosen by the agent; but CP2 can trump CP1 if avoiding A was to be chosen by the agent.

Hard to imagine indeed, and even harder to believe that such cases can be instantiated given the complexity of the neural basis of choice, but perhaps such cases are logically not impossible. The question still remains how all this can prove that an agent can be responsible without the ability to do otherwise. For if the operation of CP2 can be overridden by the agent's own choice, then it seems that the mere presence of CP2 cannot do anything to undermine the agent's ability to choose otherwise. Conversely, if the presence of CP2 is sufficient to undermine the agent's ability to choose otherwise, it is hard to see how it does *not* actively prevent him from making his own choice at all. Thus, in the Mele-Robb situation, just as in the original FSC, it is impossible to find any time during the agent's deliberative process at which the agent either (a) is not able to choose otherwise or (b) 'his own choice' is not *actually* constrained and hence manipulated by the alien causal process. Let me explain this point in some more detail.

Although it is generally agreed that conscious deliberation is not required for choice—neither for the *capacity* of making a choice about some particular option, which appears to be a more general

condition of freedom and responsibility—it can facilitate our discussion if we focus on the case of those choices that are the result of conscious deliberation. As it shall become evident, the possibility of spontaneous choice—that is choice which is not preceded by such deliberation—does not affect the issue anyway.

Suppose that the deliberative process—that is the actual causal sequence underlying the agent's choice at  $t_2$ —is indeterminate. A process can be indeterminate in the following two senses. A process can be indeterminate in the sense that it is not determined *when* it is completed; this is the kind of the indeterminacy typical, for instance, to radioactive decay. Other processes are indeterminate in the sense that they might bring about different results: perhaps it is genuinely indeterminate whether or not a coin tossed will come up heads or tails. These two kinds of indeterminacy can in fact be combined: whether or not the coin will come up heads might depend on *when exactly* it lands.

A process leading to an agent's choice can be indeterminate in both of these senses. A deliberative process might end at any point; that is at any time between  $t_1$  and  $t_3$ . In many cases it is the agent's choice itself which determines when her deliberation ends; and when it ends determines how the agent chooses. If it ends at  $t_m$ , the agent chooses to do A. If it ends later at  $t_n$ , then the agent chooses to avoid doing A. Then at  $t_1$ —at the beginning of the deliberative process, let us say—the agent must still be able to choose to do A and must also be able to choose to avoid doing A; since it is not determined whether or not the process ends at  $t_m$  or at  $t_n$ .

As we have seen, for the Mele-Robb case to work, we need to assume that the agent's deliberative process is indeterminate. However, if it is indeterminate in the sense that it is not determined when exactly it terminates, then the Mele-Robb scenario must be much more complex than 'standard' trumping preemption cases are supposed to be. If *only one* determinate process were launched which would cause the agent to choose\* to do A at  $t_2$  unless her own indeterminate deliberative process terminates by that choice at that time, the agent could still have the ability to choose to avoid A at any time *earlier* than  $t_2$ . For *one determinate process* launched at  $t_1$  can only terminate at *one determinate time* ( $t_2$ , let us say); otherwise, as we have just seen, the process must in a sense be indeterminate.



Thus, in order to prevent the possibility that the agent chooses to avoid doing A at a time earlier than  $t_2$ , we need to introduce *many*—in fact, given that time is continuous, infinitely many—alternative deterministic sequences. And only all of them operating together can guarantee that the agent would choose (or choose\*) to do A at any time between  $t_1$  and  $t_3$ . Moreover, if the agent has not yet made his choice at  $t_n$  (i.e. any time before  $t_3$ ), then his *not making any choice then* should also abort the completion of the alternative process just as his making the choice to do A would. Thus, the case involves a massive amount of a rather strange sort of qualitative overdetermination.

But the real problem is not its quantity. The real problem is how this massive overdetermination can obtain without constraining the agent's capacity to choose in a way that is incompatible with his freedom. In fact, it is contentious how this can happen at all without undermining the argument's assumption that deliberation is an indeterministic process. Libertarian critics argue that this scenario does not describe a form of potential intervention which can guarantee that a certain option is going to be chosen by a certain time without actually bringing the choice\* about. For this sort of causal structure seems to *actually* constrain the agent's choice in a way which is incompatible with her freedom and responsibility.<sup>16</sup>

This libertarian worry seems certainly justified, to the extent that we interpret indeterminacy in the way we have done so far. For it seems that the presence of the many deterministic processes which are needed to guarantee that the agent will choose to do A undermines the required indeterminacy of the deliberative process itself. The process can remain indeterminate in the sense that *when* it ends depends on the operation of the agent's indeterminate deliberative process. But as we have seen, this is not the relevant sense of indeterminacy in this context. We want the process to be indeterminate in the sense that it is not determined until a certain time  $t_3$  whether the agent will choose to do A or will choose to avoid doing it. But *there is simply no such time* during the deliberative process if the alternative causal process can 'block' the agent's choosing to do A *at any instant* earlier than  $t_3$ .

Suppose, however, that the agent *must* make his choice exactly at

<sup>16</sup> See Kane 2003.

$t_3$ ; i.e. suppose that the deliberative process can be indeterminate in the sense that although it must end at  $t_3$ , it is not determined *how* it ends or *which result* it produces until  $t_3$ . This seems a weird assumption about how a deliberative process operates, but let us bear with it. This sort of indeterminacy implies that, at any time  $t_n$  earlier than  $t_3$ , the agent must still be able *both* to choose to do A at  $t_3$  and to choose to avoid doing A at  $t_3$ . Without this assumption we cannot make sense of the idea that the deliberative process resulting in the agent's own choice is indeterminate. For if at any time earlier than  $t_3$  the agent is *already* unable to choose to avoid doing A at  $t_3$ , then the process of deliberation is obviously determinate; at least in the sense which is relevant to FSCs. Of course, the process might be indeterminate in some other sense; it can, for instance, be indeterminate which physical mechanism realizes the agent's deliberative process. But such kind of indeterminacy is obviously irrelevant from the deliberative perspective.

However, if the agent must be able to avoid choosing to do A at any time earlier than  $t_3$ , this poses a serious challenge to the Mele-Robb case, and in fact to FSCs in general. According to that challenge—emphasized mostly, but not exclusively, by some compatibilists—S in FSCs can be responsible only because a *merely potential* intervention cannot undermine an agent's ability to act otherwise.<sup>17</sup> Consider the time between  $t_1$  and  $t_3$  when the agent has not yet made her choice whether or not to do A. By assumption, his process of deliberation right after  $t_1$  (at  $t_2$ , let us say) must still be indeterminate in the sense that, at that time, she is still able both to choose to do A *and* to choose to avoid doing A at  $t_3$ .

But this means that at time  $t_2$  the presence of the alternative deterministic sequence, which *would* at time  $t_3$  override the agent's choice to avoid doing A if she was to make that choice then, cannot deprive her of the ability to choose so at  $t_3$ . Even if the alternative causal process is in progress, the agent must be able to choose to avoid doing A at  $t_3$  earlier at  $t_2$ ; otherwise the process leading to the choice of A could not be indeterminate. And exactly the same will

<sup>17</sup> For different versions of this response to FSC see Vihvelin 2000 and 2004, Smith 2003, Fara 2008, Huoranszki 2011, Nelkin 2011. For an early incompatibilist reply to FSC based on the same idea see Lamb 1993.

hold at any instant between  $t_2$  and  $t_3$  during the deliberative process. However, and crucially, the agent's situation at  $t_2$  *does not differ in any relevant respect* from the situation in which she is before  $t_2$ . Given that she chooses to do A, the alternative process remains uncompleted and hence totally irrelevant from the point of view of the agent's ability to choose to avoid doing A at  $t_3$ .

Consequently, we have no reason to assume that just because an agent *would become* unable to choose to do something in some *alternative circumstances* in which she *does not do* what she actually does, she cannot *actually* have that ability to choose to act otherwise either. For observe that the *deterministic* causal sequence CP2, which would terminate in causing the agent to choose to do A, *exists merely counterfactually*, not in actuality. If the agent's deliberative process terminates by the agent's choosing to do A, then simply *there is no such deterministic actual causal sequence* that would cause her choice to do A. Mind you: a causal *process* can be deterministic *only* in the circumstances in which it *does* terminate by causing an effect E at  $t_2$  if it is launched at  $t_1$ . If a causal process is launched and it *may or may not eventually cause* E, then it can only be an *indeterminate process*; a process the causal completion of which *depends* on what else happens indeterministically in its causally relevant environment.

This means that the agent who at  $t_3$  chooses to do A at  $t_4$  is in exactly the same situation as he was earlier: he has not lost his ability to choose to avoid doing A. In fact, if he chooses at  $t_3$  to do A at  $t_4$ , then there is no such alternative *deterministic* causal process that could cause him to choose doing A. Moreover, whether or not there is any causal sequence which could cause him to choose to do A depends on how *he* chooses. For if he does choose to do A, then the *deterministic* causal process which *could* deprive him of the ability to choose to avoid doing A does not exist. Thus, he is able both to choose to do A and to choose to avoid doing A; and it is at least partly because of having this ability that we think he is free and responsible for what he does.

Of course, one could still insist that the mere presence of CP2 can deprive the agent of her ability to choose to act otherwise at  $t_3$ . However, anyone who so insists must also admit, by parity of reasoning, that the agent is also unable to choose to avoid doing A *at any time* earlier than  $t_3$ . And if the agent is unable to choose to avoid doing A

at any time between  $t_1$  and  $t_3$ , then *the process of deliberation* cannot be indeterminate in any conceivable sense. That is if the alternative (potential) deterministic sequence is considered as *already* making the agent unable to choose to avoid doing A any time before  $t_3$ , then the mere presence of that sequence must be interpreted as *an actually manipulating factor*, which does undermine S's responsibility for his action. But so interpreted, FSCs cannot prove the irrelevance of alternative possibilities either. One can hold on to the 'intuition' that the mere presence of an alternative, preempted or trumped causal process can actually deprive an agent of his ability to choose any another option than what he actually does *at any time* before his choice has been made. But if this is indeed one's 'intuition', then one cannot at the same time hold that the agent in FSCs is free and responsible since then his choice must have been *actually* brought about by such conditions.<sup>18</sup>

Thus FSCs have no *independent* force to prove the irrelevance of alternative possibilities. For *either* we interpret the mere possibility of intervention as a factor which actually undermines an agent's ability to choose and do otherwise, but then there is no reason to believe that the agent's choice is free and that she is responsible; *or* we consider the mere possibility of an intervention as insufficient to deprive an agent of her ability to choose and act otherwise, but then it is—at least partly—due to her having that ability that we hold her responsible. *Tertium non datur.*

#### 4 FSCs and compatibilism

Many philosophers think that if any of the FSCs succeeded, the traditional metaphysical debate between compatibilists and incompatibilists would lose its point. After all, that debate was over the issue whether or not agents' ability to do otherwise is compatible with the truth of physical determinism. Both compatibilists and incompatibilists assumed that the ability to do otherwise is necessary for agents' (metaphysical) freedom. But if having alternative possibilities is not a condition of agents' freedom and/or their responsibility, then even if the truth of determinism can somehow undermine the possibil-

<sup>18</sup> This appears to be Ginet's position, see Ginet 1996.

ity of freedom, it does not do so because it is incompatible with the ability to do otherwise. For FSCs show that agents can be free and responsible even if they lack that ability. In fact, this seems to have been Frankfurt's own conclusion about the significance of his cases (Frankfurt 1988: 10).

But the success of FSCs would not have any straightforward consequence regarding the compatibility of determinism with agent's metaphysical freedom. In fact, its best version—the Mele-Robb case—grants the existence of locally indeterministic processes. For, by assumption, the process that realizes an agent's deliberation leading to her choice must be nondeterministic. Hence their example presupposes either that the universe is indeterministic or that local physical processes can be indeterministic even if global physical determinism holds.

Moreover, even if an agent can be responsible without being able to choose and act otherwise *whenever* his action occurs in the circumstances of potential overdetermination characterized by FSCs, such circumstances obtain rarely, if ever. But agents' actions understood from the microphysical perspective—if it makes sense to talk about actions from that perspective at all—are either *always* determined or *always* undetermined, pending on the physical laws of our actual world. Thus it might seem that even if an agent's choices *do* contribute to the production of her own actions—i.e. her choices are not epiphenomenal but causally necessary for her physical actions—her behavior is always *actually*, not only potentially, overdetermined. Hence the challenge of physical determinism to agents' freedom is not answered by the success of FSC. For, contrary to Frankfurt's claim, if physical determinism can indeed undermine the agents' ability to act otherwise, then even if the agent does contribute in some way to the production of her own behavior, that 'contribution' seems irrelevant; the 'real causes' of behavior are always 'alien physical forces' that are 'beyond the agent's control' rather than the agent's own wants, intentions, choices or decisions.

However, and much more interestingly, some examples that are structurally similar to FSCs can explain why the existence of deterministic causal processes need not deprive agents of their ability to act otherwise. For as I shall argue next, such processes can support, rather than undermine, agents' abilities to choose and do otherwise

in a deterministic universe.

Recall the original Frankfurt-story. In that example, an agent (or process) *M* can manipulate *S*'s behavior in the sense that if *S* were to avoid choosing *A*, *M* would intervene and cause *S* to choose—or rather: to choose\*—to do *A*. But if such a situation is conceivable and causally possible, so is the situation in which there is another agent in the story—let us call her *MM*—, who wants to guarantee that *M* is *not* going to choose to interfere into *S*'s choices. Then whenever *S* is about to avoid choosing *A*, and hence *M* is about to choose to intervene, *MM* intervenes and thereby blocks *M*'s intervention into *S*'s choice. In this situation *MM* shields or shelters, as it were, *S*'s ability to choose and do otherwise.

But, and this is the crucial point, if the original FSC is compatible with the truth of physical determinism, so must be our more complicated scenario. After all, all we have done was to iterate the original case. If *S* can be in the situation as described by the original Frankfurt-cases, *M* can be as well in our more complex case. Moreover, as we have seen, we need not assume that *M* is an agent; rather, *M* can be just a sort of causal process or mechanism. But the same holds, of course, about *MM*: it need not be an agent who reacts to *S*'s pre-choice states, it might simply be an 'alternative causal process'.

There is no logical limit to iterating further the cases. Adding yet another manipulator *MMM* can shelter *M*'s ability to intervene into *S*'s choices if *S* is about to avoid choosing to do *A*. But then, of course, we can introduce *MMMM*, who or which ... and so on. Such cases with complex preemption structure are clearly possible and their possibility shows that what matters for an agent's ability to act otherwise is not the truth or falsity of physical determinism, but the particular causal structure of the local situation in which an agent chooses and acts.

But they can show even more. Suppose that FSCs are possible if physical determinism is true. This is a trivial assumption if FSCs aim to prove anything about compatibilism. If the truth of physical determinism excludes the possibility of FSCs, then the possibility of FSCs would support incompatibilism rather than compatibilism, which seems to be a paradoxical result. FSCs have been introduced, after all, to remove one important obstacle to compatibilist theories of freedom and responsibility. If FSCs were successful, this would

prove not only that alternative possibilities are irrelevant for agents' freedom and responsibility, but also that it is irrelevant whether or not the truth of physical determinism entails agents' inability to do otherwise.

However, the possibility of such more complex FSCs prove something else. What they show is that an agent's ability to act otherwise depends on the *complex causal structure of the situation* in which she chooses and acts. As we have seen, whether or not S (in FSC) has the ability to choose and act otherwise is a contentious matter. But it is clear that S in the more complex situation with both M and MM on the scene can have the ability to choose and act otherwise. And if this situation is compatible with physical determinism, then agents' ability to choose and act otherwise is also compatible with the truth of physical determinism. Hence, what the success of Frankfurt-scenarios would prove about compatibilism is not so much the irrelevance of alternative possibilities, but that determinism is compatible with such causal structures that actually *support* the agent's ability to do otherwise. Hence, if successful FSCs could be construed, they would provide as much support for classical compatibilism as for semicompatibilism.

I am not claiming, of course, that FSCs in fact do this, because I doubt that in the cases as described the potentially manipulated agents lack the ability to choose and act otherwise. Nonetheless, if the conceivability of FSCs does show anything, it shows the possibility of such causal structures that can guarantee agents' ability to do otherwise even if determinism is true. Thus, paradoxically perhaps, if a successful FSC could be used to argue against the relevance of alternative possibilities to agents' freedom, it can also be used as an argument for traditional compatibilism.<sup>19</sup>

Ferenc Huoranszki  
Central European University  
Philosophy Department  
Budapest, Nádor u. 9., Hungary, H-1051  
huoransz@ceu.edu

<sup>19</sup> I wish to thank the participants of the *II Blasco Disputatio* workshop at the University of Valencia in September of 2016 for their questions and comments on an earlier draft of this paper and especially Pablo Richter for his useful suggestions for the final version.

## References

- Dennett, D. C. 2003. *Freedom Evolves*. London: Penguin Books.
- Ehring, D. 1997. *Causation and Persistence*. Oxford: Oxford University Press.
- Fara, M. 2008. Masked abilities and compatibilism. *Mind* 117: 843–65.
- Fischer, J. M. 2012. *Deep Control. Essays on Free Will and Value*. Oxford: Oxford University Press.
- Fischer, J. M. 2002. Frankfurt-type cases and semi-compatibilism. In *The Oxford Handbook of Free Will*, ed by Robert Kane. Oxford: Oxford University Press, 2002.
- Fisher, J.M. and Ravizza, M. 1998. *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge: Cambridge University Press.
- Frankfurt, H. 1988. *The Importance of What We Care About*. Cambridge: Cambridge University Press.
- Funkhouser, E. 2009. Frankfurt cases and overdetermination. *Canadian Journal of Philosophy* 39: 341–369.
- Ginet, 1996. In defense of the principle of alternative possibilities: why I don't find Frankfurt's argument convincing. *Philosophical Perspectives* 10: 403–17.
- Hitchcock, C. 2011. Trumping and contrastive causation. *Synthese* 181: 227–240.
- Huoranszki, F. 2011. *Freedom of the Will. A Conditional Analysis*. New York: Routledge.
- Kane, R. 1996. *The Significance of Free Will*. Oxford: Oxford University Press.
- Kane, R. 2003. Responsibility, indeterminism, and Frankfurt-style cases: a reply to Mele and Robb. In *Moral Responsibility and Alternative Possibilities*, ed. by GD Widerker and M. McKenna. Aldershot, UK: Ashgate Press.
- Lamb, J. 1993. Evaluative compatibilism and the principle of alternative possibilities. *Journal of Philosophy* 90: 517–527.
- Lewis, D. 1986. *Philosophical Papers. Volume II*. New York: Oxford University Press.
- Mackie, J. 1974. *The Cement of the Universe. A Study on Causation*. Oxford: Clarendon.
- Mele, A. and D Robb. 1998. Rescuing Frankfurt-style cases. *Philosophical Review* 107: 97–112.
- Nelkin, D. 2011. *Making Sense of Freedom and Responsibility*. New York: Oxford University Press.
- Pereboom, D. 2001. *Living without Free Will*. Cambridge: Cambridge University Press.
- Schaffer, J. 2000. Trumping preemption. *Journal of Philosophy* 9: 165–81.
- Smith, Michael, 2003. Rational capacities, or: how to distinguish recklessness, weakness, and compulsion. In *Weakness of the Will and Varieties of Practical Irrationality*, ed. by Stroud and Tappolet. Oxford: Oxford University Press, 17–38.
- Van Inwagen, P. 1978. Ability and responsibility. *Philosophical Review* 87: 201–224.
- Vihvelin, K. 2000. Freedom, foreknowledge, and the principle of alternate



- possibilities. *Canadian Journal of Philosophy* 8: 1–24.
- Vihvelin, K. 2004. Free will demystified: a dispositional account. *Philosophical Topics* 32: 427–50.
- Widerker, D. 1995. Libertarianism and Frankfurt's attack on the principle of alternative possibilities. *Philosophical Review* 104: 247–61.