Alternative Possibilities and the Failure of the Counterexample Strategy

Michael S. McKenna

1. In *The Metaphysics of Free Will: An Essay on Control* John Martin Fischer bases a position he calls "semi-compatibilism" upon a refutation of the claim that moral responsibility requires the existence of alternative possibilities. Consider the Principle of Alternative Possibilities, a principle which expresses an alleged necessary condition of moral responsibility for acts:

\[ \text{PAP: a person is morally responsible for what she has done only if she could have done otherwise.} \]

It is PAP, or some cousin, that makes trouble for traditional attempts to show that moral responsibility is compatible with determinism; for if causal determinism is incompatible with the freedom of the will, then, if determinism is true, no one ever has the power to choose or act differently than one does. If no one ever has the power to choose or act differently, then no one ever has any genuine alternatives. Consequently PAP is never satisfied and therefore no one is ever morally responsible for one's actions. The majority of *The Metaphysics of Free Will* is devoted to a sustained defense of the incompatibility of causal determinism and the ability to choose or will differently than one does. However, in chapter 7 Fischer turns to a defense of Harry Frankfurt's well-known counterexamples to the condition of alternative possibilities. Fischer and others have reasonably assumed that if the counterexample strategy is successful it can be shown that alternative possibilities are not necessary for moral responsibility. This opens logical space for a split between the traditional problem of the freedom of the will, conceived as the ability to will differently than one does will, and the attendant problem of moral responsibility. On Fischer's semi-compatibilist approach, it is possible that determinism is incompatible with the ability to will otherwise, and yet compatible with moral responsibility.

I have no objections to Fischer's excellent defense of the basic incompatibilist account of the relation between determinism and the freedom to will otherwise. Nor do I wish to object to Fischer's general semi-compatibilist position. I too am concerned to defend some form of semi-compatibilism. However, I believe that Fischer's defense of the Frankfurt counterexample strategy is not decisive. Drawing on Fischer's discussion I hope to show that the advocates and the opponents of alternative possibili-
ties rely upon differing notions of the kind of control which intuitively grounds a principle like PAP. Thus, the basis Fischer and others have used to reject PAP turns on a conception of control which the incompatibilist is not directly concerned to defend. And more important, it remains unclear that there are counterexamples to the kind of alternative possibilities that the advocates of PAP have in mind.

2. We can begin with Harry Frankfurt's famous counterexamples to PAP. Frankfurt has attempted to produce a kind of counterexample in which a person might be responsible for what she does but yet it is false that she had any genuine alternatives available to her such that she could have avoided performing that action. The ingredients needed to produce a Frankfurt-style counterexample to PAP are an unactualized causal chain designed to control your behavior, an intervener who can actualize this chain, and a piece of behavior by you. Should you indicate that you will not do what the intervener wants, she then initiates the otherwise unactualized causal sequence, bringing it about that you do what it is she wishes. But in all of these examples the intervener never intervenes because you, on your own, do what the intervener wants you to do. In a case such as this, according to Frankfurt, though you could not have avoided doing what you did, since, in the actual sequence, you did what you did because you really wanted to do it, you are responsible for what you did.

Consider the following example. Suppose that a few seconds ago I jumped into the air and that I did so because I wanted to do it. But suppose that unbeknownst to me a group of neurologists implanted in my brain a mechanism which monitored my thought processes such that, were I to have formed the intention not to jump into the air, the neurologists would have used the mechanism to cause me to jump into the air. As it happened, I jumped into the air on my own and it seems that I can be held responsible for having done so; but had I intended not to, the neurologists would have performed their crafty little trick and I would have jumped. Here I could not have avoided the situation by bringing about some alternative in which I avoided jumping into the air. Thus, though I might be responsible for having done so, I could not have done otherwise, PAP is not satisfied, and therefore PAP is not a necessary condition for moral responsibility. Call this case Jump 1.

The advocate of PAP will respond in something like the following manner. She will argue that Frankfurt has failed to distinguish between tokens and types. In the contrary to fact situation in which the agent was caused by the intervener to do "the same thing," the agent actually did do something different than what she did in the circumstance in which she brought about the action of her own accord. In the particular case at hand we can say that though I could not have avoided the type occurrence of my jumping into the air, I could have avoided the particular token occurrence in which I bring it about that I jump into the air. Thus, the proponent to PAP
will say, in all the alleged Frankfurt-style counterexamples to PAP, the agent is no doubt responsible. But, she will argue, the reason that she is responsible is because there does exist something which she could have avoided in the contrary to fact situation. Fischer has called this the "flicker of freedom" strategy. The general strategy comes to the following. For each putative counterexample to PAP, the advocate of PAP will argue that indeed there is something which the agent could not have brought about (the act-type); however in each case there exists something —some flicker of freedom—which the agent could have avoided (the act-token), and that is what explains why we are willing to hold a person responsible in these kinds of cases. Therefore, PAP is not falsified.

3. In The Metaphysics of Free Will Fischer responds in two stages. In the first stage he argues that the Principle of Alternative Possibilities tells us that we are responsible for what we do only if we have some alternative available to us such that we could have avoided performing a type of action. That is, what intuitively seems to ground PAP is that we have a kind of control in which, if we are responsible for some action, then we must be able to avoid the bringing about of an action of that type. There must be an alternative sufficiently robust that we could bring it about that that type of action is avoided and some other type of action is brought about. In this way, as Fischer puts it, we exercise a robust kind of control over the future; we are able to bring about different types of future scenarios. Thus, what PAP says in Jump 1 is that I am only responsible for jumping into the air if I have an alternative such that I actually can bring it about that I do not jump into the air, but instead, say, break into a little soft shoe routine. The disputed examples do show that this kind of alternative was not available and yet the agent is responsible. It is this kind of control which the condition of free will is alleged to offer and it is this kind of control which the modified compatibilist is arguing we can do without.

According to Fischer, the advocate of alternative possibilities wrongly assumes that the only way to undermine PAP is by showing that there are no alternatives at all which are available to an agent when she acts responsibly. However, Fischer argues, one can grant that in the counterexamples there exists some alternative to what transpires in the actual sequence. What undermines PAP here is that the existence of the available alternatives cannot provide the agent with the kind of robust control which can reasonably be expected to ground our attributions of responsibility. Thus, granted that in the counterexamples, there does exist some kind of control, some flicker of freedom, it is not in virtue of this kind of flicker that we hold the agent responsible for what she does do in the actual sequence.

Here I believe that Fischer has not fully addressed what motivates the advocate of PAP, and particularly the incompatibilist-libertarian advocate of PAP. She might respond as follows. Frankfurt, Fischer, and company have gotten wrong what intuitively drives PAP. They think that the kind of
control demanded by PAP is the kind of control in which we have it within our power to bring about some range of different kinds of alternatives—like jumping up into the air, or instead doing a little soft shoe; but what intuitively drives PAP is the kind of control needed in order for us to avoid being the author of a particular act and thus avoid being responsible for the production of that particular action—like either being the one who brings it about that I jump into the air or not being the one who brings it about that I jump into the air. It is a matter of holding people accountable for what they do only if they can avoid any blame or punishment that might fall upon them for performing those very particular actions which they do perform. In all of the putative counterexamples to PAP, this kind of control is available to the agent when she acts. Should the contrary to fact intervener in Jump 1 kick in and cause me to jump into the air, it would do so because the alternative was available for me to avoid the situation in which I bring it about that I jump into the air. I may be responsible for jumping into the air, but this is because I could have avoided the situation in which I bring it about that I jump into the air. And this is a genuine alternative that was available to me even if I could not bring it about that I do some other type of thing than jump into the air.

Fischer has yet another card to play in the first stage of his response. He will object to such a narrow flicker of freedom as the intuitive basis for responsibility, not just because the alternative found in the flicker is so minuscule, but because of the kind of alternative that one finds there. For in the counterfactual scenarios the agent has an alternative available to her in which she acts unfreely. How can the existence of an alternative in which an agent does not act freely provide the intuitive support for why we ought to hold her responsible in the actual scenario in which she does what she intends of her own accord? The robust notion of alternative possibilities for PAP is understood to provide entirely different kinds of future scenarios. This explains why, when an agent chooses one pathway, she exercises a kind of control which would ground judgments of responsibility for the path she chose. But a flicker offering the control of choosing between one particular will, and a state of coercive influence gives no justification for our judgment of responsibility. It is not just that the flicker here is small, but that it offers only an alternative which itself implies a lack of control. How can the existence of a situation in which one loses control ground our account of the kind of control requisite for one's responsibility?

This way of posing the question brings into stark relief the incompatibilist intuitions as against the compatibilist intuitions which Fischer is relying upon. Fischer's account of why this kind of control fails focuses on the alternatives by projecting from the agent's will at the time of willing out onto the alternatives available. And no doubt, in these kinds of cases, the alternatives are pretty dismal. But the incompatibilist alternative is to look, so to speak, from behind the agent's will prior to the time of willing out onto the alternatives available at the time of willing. And from this point of view, the
alternatives are understood as different particular wills which an agent might have with different kinds of qualitative features, sometimes of an immensely important moral significance. The issue for incompatibilists here is whether the will which displays these features places my stamp upon the world, and whether it is up to me or not to have that particular stamp or some other as my mark upon the world. In the Frankfurt-type cases the alternatives are, either doing what one does of one’s own intention, or being coerced into performing the same kind of action against one’s will. These alternatives do seem to be quite impoverished; however, they mean all the difference between one’s doing something of one’s own will, and one’s not doing that kind of thing of one’s own will—of that kind of action not being something that comes from one’s self. Surely this is precisely the locus of the incompatibilist’s concern: an agent is responsible for the will which she has only if she has available to her power to control that very will. What more fundamental kind of control can there be here other than the control for one to either have a particular will or not have it?

4. Here is the second stage in Fischer’s response. Suppose that the advocate of alternative possibilities holds steadfastly to the existence of that small flicker, as in the above discussion I suggested that she might. Fischer concedes that it is in the nature of the counterexample strategy that there must be some triggering event which serves as the initiation of the counterfactual intervener. Whatever event that is, it is always open to the incompatibilist to insist that the flicker of freedom resides in the ability to exercise or not exercise that particular triggering event. But, up until this point, those triggering events are best understood as things under an agent’s control, like forming an intention to X, which would then cause the counterfactual intervener to cause the individual to Y. As Fischer argues, all of the counterexamples can simply be rewritten so as to place the triggering event prior to the execution of anything that might even remotely be described as within an agent’s control. Suppose, for instance, that just before a person formed any intention X which was either incompatible with Ying, or simply formed the intention to not Y, she would inexplicably though reliably blush red. If the counterfactual intervener conspires to see to it that she Y, then, should she blush red, the intervener intervenes and insures that she Y. Here the agent does not even have control sufficient to have an alternative to the intention Y. The triggering event is now a blushing, and hence the only alternative available here is to either blush or not blush. Surely, Fischer argues, no one would want to ground ascriptions of responsibility for action on the ability to either blush or not blush, since, quite obviously, such an episode is not something over which a person could normally exercise any kind of control.

Consider, then, the amended example, Jump 2: On the basis of my blushing red the counterfactual intervener can reliably predict the formation of any of my future intentions incompatible with jumping into the air, or the
formation of my future intention to not jump into the air. Should I indicate by blushing that I will not form the intention to jump into the air, the counterfactual intervener will initiate the mechanism, causing me to jump. As things go, I choose on my own to jump into the air. Since I did choose on my own, it seems that I can be held morally accountable for my jumping, despite the fact that I could not have formed even an alternative intention to the one which eventually led to my jumping.\footnote{6}

Fischer maintains that cases like Jump 2 show that in the Frankfurt-type cases, our ascriptions of responsibility cannot be grounded upon the existence of alternative possibilities. But, Fischer explains, it is open to the advocate of alternative possibilities to argue that there is some other factor which is essential for grounding our ascriptions of responsibility and that this factor requires alternative possibilities. As Fischer presents the case, the advocate of alternative possibilities faces the burden of showing what other condition might be required for grounding our responsibility ascriptions, and why it is that this factor requires alternative possibilities.\footnote{7} In what follows, I would like to consider whether the advocate of alternative possibilities can respond more directly to Fischer by arguing that alternative possibilities do ground our ascriptions of responsibility in the Frankfurt-type cases. What, then, can the advocate of alternative possibilities say about cases like Jump 2?

There are a number of puzzling difficulties which present themselves here. First, this suggestion seems to assume that there is a causally deterministic relation between the triggering event and the deliberative processes an agent undertakes. But, of course, it is precisely here where the libertarian will locate the indeterminacy required for freedom of the will. Second, these counterexample strategies also assume fairly implausible views of the mental. In this case we do not merely have the extraordinary sci-fi examples of neurosurgeons monitoring intentions, and so on; we have physical states prior to the formation of intentions used as reliable indicators of upcoming intentions, deliberations, and the like. If Davidson's views about the holism of the mental are at all convincing, there is good reason to hold that there are no psycho-physical laws.\footnote{8} It would seem that any connections between blushing and intentions would be out of the question.

However, I do not think that the advocate of alternative possibilities needs to resort to these objections. Focus upon PAP as it is construed by the incompatibilist and libertarian. What it means to say that an agent is morally responsible for the action she does perform only if she could have avoided that action, is that she is morally responsible for performing that very token action which she did perform only if she could have avoided the performance of that very action. In the contrary to fact situation in which an intervener intervenes, either as the consequence of a blushing red or the formation of an intention, it is simply false that the agent does the very same thing and that she cannot avoid the performance of that very same thing. For in the counterfactual case, her act, or more accurately, her bodily
movement, is simply not in the relevant respects the very same act token at all. It is true that in the amended counterexample strategy, the counterfactual intervention precludes the formation of any intention other than the one which the agent does form. Thus, she does not have any other intention available to her; but this does not mean that she cannot avoid doing exactly what she does in the actual sequence. For in the alternative sequence she does not do the very same particular thing. Thus the amended counterexamples cannot show that she cannot avoid doing exactly what she does. For the examples do not in fact show that she does do the very same thing.

What I believe is happening here is that the counterexamples are being pressed into doing two different pieces of work. They are supposed to show that an agent does not have it within her power to bring about a different alternative than the alternative which she does bring about. But they are also supposed to show that she could not have avoided doing exactly what it is that, in the actual sequence, she is being held responsible for doing. But, the libertarian might argue, in these amended versions, the counterexamples only show the former of these two. The libertarian advocate of alternative possibilities is going to insist that both conditions be met in order to decisively refute principles like PAP. Thus, the libertarian might grant here for the sake of argument that the agent in such a case did not have an alternative available to her in which she had it within her power to avoid doing anything other than what she did; however, it does not follow that she could not have avoided the particular thing that she did do. For the existence of a possibility in which she did not do that thing at all can hardly be used to prove that she could not have avoided it.

It will be insisted here that the incompatibilist or libertarian is employing an essentialist principle of act-token individuation and thereby trivializing the condition expressed by PAP. If the identity conditions for a particular act-token are such that any variation in the situation related to that action will result in a numerically distinct act, then obviously the advocate of PAP gets her way. But on this strict version, a person exercises her ability to act differently than she does (in the sense relevant for moral responsibility) every time she alters the trajectory of the wave of her hand in the air, or the particular fork she chooses to draw from the utensil drawer. However, this is not required. All that is needed in order to insulate against the counterexample strategy is an intuitively clear understanding that whatever the identity conditions might be for the same act-token (or act-particular) there will be certain salient features necessary for sameness which will certainly be ruled out by nonstandard sorts of causal interventions in the normal deliberative machinery of rational agents. This will allow sufficient leeway in our counterfactual analysis of "the same act-token" so as to insure that there is substance to claims of identity. Thus, in some other possible world, should I pick the butter knife from the middle of the pile of knives in the utensil drawer, I would have performed "the same act-token"
of "fetching a knife for the butter" that I did perform in this actual world in
which I fetched the knife from the right hand side of the pile. However,
should I be caused in some other possible world by a neurological manipu-
lator to go to the utensil drawer to get a knife (perhaps the very same knife
which I pick up in this actual world), this would be a very different act-
token than the one which is caused by the proceedings of my mundane
little thoughts about breakfast, knives, and butter.

5. Suppose now that the incompatibilist and libertarian have an ade-
quate account of act individuation and that the case can be made that the
revised examples do not show that the agent does the very same thing in
the alternative sequence. Fischer might now respond by appeal to consid-
erations previously employed to handle the original incompatibilist objec-
tions. Here again, Fischer might argue, the advocate of alternative possi-
bilities assumes that the only way to undermine PAP is by a demonstrable
proof that there are no possibilities available to the agent other than the one
which transpires in the actual sequence. But, again, the examples will effec-
tively undermine PAP if they undermine the kind of control which puta-
tively grounds our ascription of responsibility in the actual sequence. Be-
cause the amended examples like Jump 2 preempt the ability to form any
intention other than the intention which the agent does act upon, the ex-
amples do clearly undermine the agent's power to have any will other than
the particular will she does have. Thus, strictly speaking, in Jump 2, in the
alternative sequence the agent does not perform an act numerically identi-
cal with the one she performs in the actual sequence. There are therefore
alternatives to the actual sequence. However, the alternatives in Jump 2 do
not provide the agent with even the flicker of control over her own inten-
tions. Since the agent has no control over this kind of alternative, it is not an
alternative that is accessible to her. Numerically identical or not, it cannot be
in virtue of this kind of alternative possibilities, one completely inaccessible
to the agent, that we justify our ascription of responsibility in the actual
sequence.

The advocate of PAP will be unwilling to concede that the amended
counterexamples do genuinely show that the agent did not have the power
or ability to have any intention other than the intention which she did have.
The incompatibilist and, most notably, the libertarian want to locate the
freedom relevant condition for responsibility in an intrinsically important
feature of the will, and at this point the libertarian will want to distinguish
between having the power to have a different intention (or will) than what
one does have, and the access to a particular alternative intention (or will)
on some particular occasion.

The libertarian can make this point by drawing a distinction between
the ability to exercise a power, and the having of that power. What the
amended examples do is preempt the ability to exercise one's power in some
counterfactual scenario; but they do not show that in the actual sequence the
individual actually fails to have that power to have a different will. To see this, consider what transpires, according to the libertarian, in the actual sequence of one of these amended counterexamples. At a certain time, time t, S forms the intention to C. In the actual sequence of events, say at time t-1, S exercised the following power: to form either the intention C or the intention ~C. As it happened, in the actual scenario, S formed the intention C at time t. Now in the actual sequence as it transpired, the libertarian will insist that if this kind of power at t-1 was not genuinely available to the agent, then she simply cannot be held accountable for what she has done at t or at t+n. For the libertarian it is crucial that in the actual sequence of events leading up to the formation of the intention to C, that that sequence include the point at which the agent exercises this deliberative capacity to intend to C, or ~C. Because in the actual scenario, the individual exercises this power in the manner preferred by the counterfactual intervener, the individual is left alone to exercise this power as she sees fit. In the counterfactual scenario, it is true, at, say time t-2 that the agent is preempted from forming the intention to ~C. But on the libertarian view of the matter, this is only because the triggering mechanism (say a blushing red) is sensitive to how the individual would have exercised her power at t-1. Thus, the intervention at t-2 precludes the agent's ability to exercise her power at t-1, but it is only in virtue of being sensitive to how the agent would otherwise have exercised this very power at t-1 that, at t-2, the triggering mechanism is initiated.

The crucial point here is that the libertarian will include in her account of the actual sequence, the exercising of a power to intend, choose, or will otherwise. It is of no consequence that in the counterfactual scenario, an agent is preempted from exercising this capacity in any manner other than the way she exercises it in the actual sequence; what is relevant to an analysis of the actual sequence is that particular power which essentially involves alternatives. What drives the libertarian view of control, then, is a picture of agency in which we are the ultimate sources of our intentions and in which it is up to us whether we have the particular intentions, choices, deliberations, and wills which we do have. Having the power entails having the alternatives. All counterfactual intervention can do is preempt the ability to exercise a power of this kind, but it does not mean that, in an analysis of the actual sequence of events leading to actions, there is no such power available to the agent. The libertarian's position is based upon the intuition that an individual must be able to avoid having the very particular intentions, choices, wills, and so forth which she does have in order for her to be held responsible, and to be subject to praise and blame. What amended counterexamples like Jump 2 show is that in the alternative sequence an agent might be precluded from exercising this kind of power, but it does not mean that, in the actual sequence, this kind of power was not possessed and exercised by her.

6. Fischer might object at this point to the libertarian's insistence that
the agent does in fact have a power in the actual sequence which can accurately be characterized as the power to intend to C or ~C. He might set out the following principle as a general principle required for an agent to have a power:

PP: If S has the power in circumstances A to C (or ~C), then S can C (or ~C) in circumstances A.  

Call this the Principle of Power.

In Jump 2, Fischer will argue that PP shows that the libertarians' account of control in the actual sequence is simply inaccurate. To characterize the agent's power in the actual sequence as the power to intend to C or ~C is simply to misrepresent what transpires. If there is to be substance to the claim that S can genuinely intend to C or ~C in Jump 2, then one must show that S can intend to ~C. But the counterfactual intervener insures that half of the disjunct cannot be satisfied. S simply cannot intend to ~C. Therefore in the actual circumstances A, it is false (or trivial) that S has the power to intend to C or ~C. S simply has the power to intend to C.

How might the libertarian respond? She might demand clarification of what is involved in the circumstances mentioned in PP. If circumstances A are fixed by the actual scenario, including any kinds of counterfactual factors or overdetermining circumstances, then perhaps in cases like Jump 2, PP shows that S does not exercise a power to intend to C or ~C. But if circumstances A in PP are fixed, not by the actual scenario, but by the actual sequence which brought about S's jumping, then in capturing the relevant kind of libertarian control to intend to C or ~C, the libertarian can reasonably argue that the counterfactual intervention would have to be factored out of the circumstances A relevant to PP.

The libertarian is attempting to locate a power of responsible agents in the deliberative mechanism which results in action. Consider the point t-1 in Jump 2 as it occurs in the actual sequence. The libertarian insists at t-1 upon an agent's exercising of a power characterized as the power to intend to C or ~C. Now, what are the circumstances at t-1? The counterfactual mechanism is set up to predict reliably what will transpire at t without the aid of intervention. Should the mechanism predict that S will ~C at t, then, at t-2 the mechanism preempts the formation of the intention at t-1. But according to the libertarian, in the actual sequence, at t-1, S exercises a genuine power to intend to C or ~C. In the actual sequence, the counterfactual mechanism at t-1 is not operative. The fact that the only available outcome in the actual scenario is that S intend to C, does not mean that at t-1 in the actual sequence S did not have the power to C or ~C. For in the actual sequence, the fact that S chose to exercise that power at t-1 by intending to C at t, does not mean that at t-1, S could not have instead intended to ~C. If S does not exercise her power in a certain way at t-1 this does not mean that she lacks the ability to do so. Thus, if the libertarian carefully specifies the circumstances in Jump 2 in the actual sequence at t-1, she can argue that S does retain the power to intend to C or ~C. This power can be characterized as follows:
At $t-1$ S has the power to intend to C or $\neg$C. S intends to C at $t$. If S had preferred to $\neg$C at $t-1$, and had no external conditions prevented S from doing so, then S would have formed the intention to $\neg$C at $t$. The libertarian can embrace PP and argue that the circumstances A require exclusion of the counterfactual element. Therefore, in Jump 2, the agent can properly be said to have the power to intend to C or $\neg$C, since it can be shown, by factoring the counterfactual intervention out of the circumstances, that the agent could intend to $\neg$C.

The incompatibilist and libertarian are concerned to locate the freedom relevant condition of responsibility in an intrinsically important capacity of the will. This treatment of the circumstances in A simply isolates the processes involved in the agent’s own deliberative capacities and places within that process a power to opt between genuinely open alternatives. To produce a counterexample here one would have to show that in the actual sequence the alternatives were not genuinely open to the agent. But in order to do this, counterfactual intervention will not work. One needs a deterministic relation which excludes the availability of alternatives within the actual sequence. This is to beg the question on behalf of determinism, and this is precisely what the libertarian wants to rule out.

It is worth noting here that there are many garden variety statements of power the analysis of which requires the elimination of extraneous variables. Take a case of overdetermination. Suppose that my car quits running altogether and that my (trustworthy) auto mechanic learns that two factors simultaneously but independently caused the car’s failure. One was that the fuel pump broke and the engine received no fuel. Another was that the electrical system burned out of the vehicle and it could not carry any charge to create a spark. Consider now the first causal explanation, that the fuel pump broke. For that explanation to show that the fuel pump was an independently sufficient cause of my car’s failure, certain counterfactuals have to be true. Most notably it has to be true that if the fuel pump had not failed, and if there were no other conditions independently sufficient for my car’s failure, then the car would not have quit running. To show here that the failure of the fuel pump was causally sufficient in bringing about my car’s failure, other causal conditions must be factored out of the counterfactuals relevant to ascertaining the truth of the causal claim under consideration. The libertarian can insist upon a similar treatment of the power relevant for ascriptions of responsibility. To understand this kind of power one must focus upon what transpires in the actual sequence of events, not in the actual scenario. For there are other causal variables in the actual scenario which might pollute a proper characterization of the power actually at work in the actual sequence.

7. What is the upshot of these considerations? The second stage in Fischer’s attack on alternative possibilities will not sway the incompatibilist and libertarian advocate of PAP. The amended counterexamples which
Fischer and others have offered simply do not show that an agent could not have avoided doing that for which she is held responsible. In the counterfactual scenario she simply does not do exactly what it is that she does in the actual sequence scenario. Furthermore, the examples do not show that, in the actual sequence, an agent fails to have control over her intentions which requires the existence of alternative possibilities; what the amended examples show is that in the counterfactual scenario, the agent is unable to exercise the kind of control which the libertarian is at pains to defend. This leaves us with Fischer's first stage of defense. Here I have argued that Fischer's case can only be made by reliance upon an intuitive view of control which the advocate of alternative possibilities does not share.

I would like to thank Peter French for his patient and generous supervision of this project. I am also grateful to the National Endowment for the Humanities for providing me with the opportunity to participate in "Responsibility and Social Issues: Theory and Practice", a 1995 NEH Summer Seminar for College Professors. The final version of this paper is the product of that seminar. I have benefited greatly from the helpful advice of David Copp, John Martin Fischer, Christopher Fitzmartin, Carl Ginet, Ishtiyaque Haji, Steven Schwartz, and George Thomas.

Bibliography


Notes


2 Fischer's target is a set of related principles all of which require the existence of some kind of alternative in order to hold an agent responsible for something. If the agent is responsible for an action, then she must have had an alternative action available to her; if it is a consequence of her doing, then she must have had available to her the ability to bring about some other consequence; if it is a failure to act, an omission, then she must have been able to do whatever it is that she failed to do. In what follows I will focus only upon PAP, a principle of action, but what I have to say can readily be applied to
any of the related principles demanding alternative possibilities.

Frankfurt's original defense of the counterexample strategy came out in his 1969 paper, "Alternate Possibilities and Moral Responsibility." Frankfurt's "What We Are Morally Responsible For" was later presented as a response to Peter van Inwagen's criticisms of the counterexample strategy in "Ability and Responsibility." Van Inwagen's criticisms of the counterexample strategy are presented in full in An Essay on Free Will.

Those familiar with the details of the counterexample strategy will see an obvious shortcoming with this formulation. The case would be more convincing if the counterfactual intervention were placed prior to the formation of any intention. I will consider a case like this later. At this point I ask for the reader's patience. Presently I am concerned to bring into relief two very different views of control and it is most economical to do so with cases like the one under discussion. Furthermore, my treatment here paces Fischer's defense of the counterexample strategy.

I purposely apply the token-type distinction to PAP as opposed to some principle about omissions or consequences. In An Essay on Free Will van Inwagen actually concedes PAP to Frankfurt. He then proceeds to argue that in the case of consequences, the distinction between particulars and universals can be adequately employed to defend alternative possibilities. I believe that van Inwagen was willing to concede PAP too easily here. He suggests that the distinction between particulars and universals cannot be carefully employed in the case of actions because he does not have available a precise notion of act-universals, and in the case of act particulars, he does not think that what we are responsible for are voluntary movements of human bodies (van Inwagen, pp.179-80). However I believe that one can quite easily make do here with intuitively clear distinctions between tokens and types.

Furthermore, one could argue that an act-token is numerically identical with itself only if it is caused by the same reasons. This specifies a merely necessary condition and is consistent with the spirit of van Inwagen's treatment of event-particulars: an event is identical with itself if and only if it has the same causes (van Inwagen, p.169). The identity condition I offer for act-tokens is also sufficiently loose so as to allow for counterfactual considerations: in the actual world I shoot the bank robber because I believed that she was carrying a hand gun; in the counterfactual world I shoot the bank robber because I believed she was carrying a machine gun; these alternative world scenarios give rise to the same reasons for my shooting the bank robber: she was carrying a gun and I believed that she was about to shoot me. On the other hand, should I shoot the bank robber because I was caused by a neurological manipulator to do so, the reasons are clearly not the same and the acts are not token-identical. (Should someone point out here that the manipulator could cause me to have the belief that the bank robber is carrying a gun, the argument can simply be pushed back a step further: this is not the same belief as the one caused in the normal kind of way.)

Fischer, pp.178-83.

Ibid., pp.140-41.

This incompatibilist or libertarian intuition is distinguished in "Two Concepts of Accountability," an unpublished paper by Gary Watson. Watson provides a convincing account of the practical social transactions and relationships which explain the intuitive source of this concept of responsibility. I believe that Watson's insight can be used to defend the kinds of intuitions which motivate the advocates of PAP.

Carl Ginet has suggested in conversation that the counterfactual intervener could manipulate by bringing it about that I bring it about that I jump into the air. This does seem to make trouble for the response I have sketched on behalf of the libertarian and incompatibilist. But again, I believe the libertarian can reasonably argue that when an agent's actions, or the deliberative machinery of her actions, are brought about by reasons independent of the agent's own rational machinery, then the actions or the deliberative machinery are not hers. This is not to say that it is a priori impossible for an intervener to cause an agent to act, or even to cause her to have such and such deliberations. It is only to say that the actions or the deliberations are not hers.

Fischer, pp.141-42.

Ibid., pp.144-45.

See Davidson's "Mental Events," and "Philosophy as Psychology," in Essays on Actions and Events.

As regards the charge that these kinds of cases assume a deterministic relation between agent and action, it might be argued that mere high probability in a nondeterministic world would be sufficient. This has been suggested to me by both John Fischer and Tim O'Connor on separate occasions. Also, Frankfurt makes a similar point in a footnote in his original paper, "Alternate Possibilities and Moral Responsibility" (p.6, n.3). However, this seems dubious since we are considering a case in which the intervener remains unactualized. Thus, in the actual world, we cannot say with certainty that the agent genuinely could not have done otherwise, though it does seem as damning to say that it was highly probable that she could not have done otherwise. David Copp has suggested just making the counterfactual intervener God who still knows the facts in an indeterministic world. This, however, does not seem all that satisfactory either since we simply replace causal determinism here with divine determinism.

As for the Davidsonian worries about the impossibility of psycho-physical laws, Ishtiyaque Haji has suggested that all one needs is a law which relates neurological states with some bodily movements. Here again, I am uncertain of the suggested response. It turns on how much one can read from the physical description of an agent's body. Even if we were to discover reliable lawlike relations between neural states antecedent to the formation of intentions, and bodily movements subsequent to the formation of intentions, it is unclear how we could discern, assuming that the lawlike statements are expressed purely in the terms of a closed physical theory, what kind of action is correlated with the state of the body at any one time.

While I remain skeptical of these suggested defenses of the refined counterexample cases, the grounds I have offered as objections to these defenses are equally specious, and I am not prepared to defend the legitimacy of the incompatibilist position on such uncertain terms.

I am especially grateful to Peter French for helping me to see this point.

Believe that it is precisely here that van Inwagen and Frankfurt are talking past each other. This explains why van Inwagen offers his account of event-particulars (van Inwagen, Essay, pp.166-70). He simply assumes that all that is needed is to show that the identity of the two alternatives are simply not the same. Thus, for van Inwagen the issue turns on whether the agent brought about the very same particular in the counterfactual sequence. Alternatively, in response Frankfurt insists in "What We are Morally Responsible For" that all that we are responsible for are our bodily movements, not the consequences of our bodily movements (p.103). He seems completely disinterested in the possibility that the universal-particular distinction, or as I have used it, the type-token distinction, can just as easily be applied to bodily movements or actions as they can to events or states of affairs. Thus, for Frankfurt the issue seems to turn upon an agent's power or ability to avoid a certain action, regardless of whether, in the counterfactual scenario, the agent does do exactly what she does in the actual sequence.

I am thankful to David Copp for helping me to get clear on what is involved in this problem.

There may still be some residual worries about other kinds of counterfactual intervention which does not look all that "nonstandard." Frankfurt uses the example of driving off the road because a person fails to keep her eyes on the road. Should she not have failed to keep her eyes on the road because the scenery was interesting, then some other preoccupation would have caused her to fail to keep her eyes on the road ("What We Are Morally Responsible For," p.101). This kind of case involves a failure to act and since I am discussing PAP, a principle not about omissions but rather about actions, it would be more useful to consider a similar case about actions. So, suppose that Jones decides to shoot his business associate Smith and does so because Jones wants to secure his place in his company. However, had Jones decided not to shoot Smith for this reason, he would have decided to shoot Smith on the basis of his belief that Smith had won his wife's affections. In the counterfactual scenario would it be correct to say that Jones
performed the same act-token that he did perform in the actual world? Perhaps here a reliance on an intuitive understanding of "same act-token" is inadequate. But I believe that the following necessary condition for act identity will be sufficient to handle these kinds of cases: an act is numerically identical with itself only if it is caused by the same reasons. I cannot defend this view here, but it does have a straightforward plausibility which recommends it, and it preserves the view of PAP which the incompatibilist is at pains to defend. With this identity condition Jones does not perform the same act in the counterfactual situation which he does perform in the actual world. It goes without saying, of course, that in either case Jones can be held responsible for what he does, but what is crucial is that he does numerically different things in each case.

All of this can be said without any assumption of trans-world identity for act-tokens. One would need a closest successor theory for act-tokens.

Fischer, pp.140-41.

I must thank Ishtiyaque Haji for his helpful advice in capturing the incompatibilist view of control which I have been developing here.

This principle, as well as the response on behalf of Fischer, was kindly suggested to me by Ishtiyaque Haji.

This counterfactual qualification is not intended as a general treatment of power. So, for instance, in Locke's famous case, the man unknowingly locked in the room lacks freedom of action though he willingly stays there. He does not have the power to leave the room. Though, if the suggested treatment of libertarian power were applied to the freedom of action, he would leave the room if he preferred and if no other external circumstances had prevented him from doing so. But this qualification is illicit. The concept of free action and the power involved in this kind of freedom can be undermined by external constraints. However, the libertarian account of the control relevant for the freedom of the will—the freedom to form alternative intentions—does not illicitly make use of this counterfactual strategy. The point of the libertarian position is precisely to locate the freedom relevant capacity for ascriptions of responsibility in an intrinsically important feature of the will. To exclude external causal variables in characterizing this power is only to isolate the power in a way that makes perspicuous how alternatives are involved in the exercising of that power.

Ishtiyaque Haji has suggested one final stage in this dialectic. Suppose that Fischer has a response to the libertarian here and can show that the libertarian conception of power cannot satisfy PP in Frankfurt-type cases. The libertarian might give up the condition of alternative possibilities altogether here, and yet argue that the counterexamples cannot be used to show that ascriptions of moral responsibility are compatible with determinism. For here, the libertarian might argue that determinism undermines ascriptions of responsibility in the actual sequence, not by eliminating alternative possibilities, but by showing that the agent is not the ultimate source of her actions. If it can be shown that an agent's actions can be in principle explained by appeal to events prior to her birth, and the laws of nature, then the libertarian might argue, the agent is not the ultimate source of her actions. On this view, what is required for responsibility is a deterministic break which allows the agent to function as the sole and ultimate source of her actions. In this case, showing that the agent has alternatives available to her might be construed as a sufficient condition for the existence of such a break, but not a necessary condition.

This incompatibilist conception of control is nicely captured in Robert Kane's "Two Kinds of Incompatibilism." There Kane specifies an "ultimacy" condition for responsible action. The ultimacy condition requires that the sole basis for the very action which an agent is responsible for must ultimately rest upon the agent. Whether or not that very action occurs turns upon the agent as its sole and final explanation (R. Kane, "Two Kinds of Incompatibilism," pp.120-21). I will not pursue this line of argument here as I believe that the libertarian has a reasonable response to the charge that PP goes unsatisfied in the amended counterexamples.

In "Alternative Possibilities and Freedom," delivered at the 1994 APA Central Division Meetings, Joseph Kiem Campbell used these insights to criticize Frankfurt's strategy. However, he attempts to draw traditional compatibilist conclusions from them.