A powerful and pervasive view holds that an individual can be morally responsible for his actions only if he is able (at some suitable time) to do other than he actually does. This is the principle of alternative possibilities (PAP). Because causal determinism appears to threaten our freedom to do otherwise, the principle is a potent tool in the argument that causal determinism rules out moral responsibility. If there are counterexamples to the principle, then at least this particular route to the incompatibility of causal determinism and moral responsibility is blocked.

Harry Frankfurt\(^1\) has presented a series of putative counterexamples to PAP. Here is his description of such a case:

Suppose someone—Black, let us say—wants Jones to perform a certain action. Black is prepared to go to considerable lengths to get his way, but he prefers to avoid showing his hand unnecessarily. So he waits until Jones is about to make up his mind what to do, and he does nothing unless it is clear to him (Black is an excellent judge of such things) that Jones is going to decide to do something other than what he wants him to do. If it does become clear that Jones is going to decide to do something else, Black takes effective steps to ensure that Jones decides to do, and that he does do, what he wants him to do. Whatever Jones’s initial preferences and inclinations, then, Black will have his way. . . . Now suppose that Black never has to show his hand because Jones, for reasons of his own, decides to perform and does perform the very action Black wants him to perform. In that case, it seems clear, Jones will bear precisely the same moral responsibility for what he does as he would have borne if Black had not been ready to take steps to ensure that he do it. . . .\(^2\)

It appears as if Jones is morally responsible for his action, although he could not have done otherwise (in virtue of Black’s presence). If this is so, then PAP is called into question (along with the argument that employs it to get to the conclusion that causal determinism is incompatible with moral responsibility).

The “Frankfurt-style” examples have evoked considerable discussion. One general form of response to the examples is that, despite the initial appearance, there really are alternative possibilities. This

---


kind of response comes in many different specific versions, but the general idea is that one can find at least some alternative possibility—some "flicker of freedom"—even in the most sophisticated "Frankfurt-style" example. Typically, the proponent of the flicker of freedom response to Frankfurt concedes that the agent cannot do otherwise (cannot perform a different action from the one he actually performs), but insists that there is something else that is the ground of moral responsibility and to which there is indeed an alternative possibility. Perhaps this "something else"—this flicker of freedom—is the agent’s performing the action "in the normal sort of way" (that is, not as a result of Black’s intervention), or the agent’s volition to perform the action, or the event particular (finely individuated) brought about by the agent in acting, and so forth. The flicker theorist thus defends something like PAP—some principle that links moral responsibilities to alternative possibilities of some sort or another.3

James W. Lamb4 gives a strikingly different response to Frankfurt: he does not even concede that the relevant agent (say, Jones) cannot perform a different action from the one he actually performs. That is, Lamb argues that in the Frankfurt-style examples there is no reason to believe that the relevant agent cannot do other than he actually does. If Lamb is correct about this, then Frankfurt and the vast majority of commentators on his work are wrong about the examples, and the examples pose no threat to PAP.

Lamb offers an “informal” argument for his view that it is a “fallacy” to interpret the Frankfurt-style examples in the “standard” way, that is, as containing an agent who cannot do otherwise. He also attributes to Frankfurt a kind of reasoning in support of the standard interpretation of the examples which he shows to be fallacious. We believe Lamb is mistaken in his reasoning in his informal

3 A careful discussion and evaluation of the flicker approach is beyond the scope of this paper. We wish simply to register our view that this approach is unsatisfying. Although we are willing to admit that one can identify certain sorts of alternative possibilities, even in the fanciest Frankfurt-style examples, it is implausible to think that these alternative possibilities play the role required by the flicker theorist; that is, it is implausible that these alternative possibilities ground the ascriptions of moral responsibility. Put slightly differently, even if there are alternative possibilities in the Frankfurt-type cases, it does not seem natural or appealing to suppose that it is in virtue of the presence of these alternatives that the agent is morally responsible for his actions. For developments of this sort of response to the flicker approach, see Fischer, “Responsibility and Control,” this JOURNAL, LXXXIX, 1 (January 1982): 24–40; and The Metaphysics of Free Will: An Essay on Control (Cambridge: Blackwell, forthcoming), ch. 7.

argument. We shall begin by explaining this mistake. Then we shall be in a good position to see why Lamb is also incorrect in attributing the fallacious form of reasoning to Frankfurt in support of the standard interpretation of the examples.

First consider Lamb’s informal explanation of the alleged mistake in the standard interpretation of Frankfurt’s examples:

Let us remove Black from the counterexample. We then have Jones “acting on his own” in such a way that he could have refrained (since otherwise Black’s presence in the counterexample would not be needed). Moreover, the triggering event does not occur. Now, restore Black to the counterexample. What has changed? I suggest only this: that it is now true that were the triggering event to occur, Jones would be forced to act. But the triggering event by hypothesis does not occur. How, then, can Black’s presence alter the fact that, in the absence of the triggering event, Jones could have done otherwise? If Jones can refrain in the absence of the triggering event when Black is not around, why can he not refrain in the absence of the triggering event when Black is present (523)?

But consider now an ordinary electrical circuit without a fuse. And suppose that it never actually is subject to a voltage surge and thus never actually overheats. Nevertheless, it is reasonable to say that the circuit could overheat and start a fire. Now imagine that we add a fuse to the circuit. Again, suppose that the circuit is never actually subject to a voltage surge. What has changed, given the installation of the fuse? It is now true (as it was not before) that, if the triggering event (a voltage surge) were to occur, the circuit would be prevented from overheating. Has anything else changed? Lamb thinks not. Indeed, his analysis of the Frankfurt examples commits him to the view that the circuit can still overheat in the absence of the voltage surge even though the fuse has been installed. But this is surely false. Once the fuse is installed, the circuit cannot overheat (even in the absence of the voltage surge). Therefore, Lamb’s parallel claim that in the absence of the triggering event Jones can do otherwise must be false.

Indeed, Lamb’s reasoning can be shown to be faulty by a whole host of examples involving the installation of a safety device. Consider, for example, the sort of “surge protector” one can buy at a hardware store to protect against power surges. Let us suppose that before you install your surge protector you do not actually have any power surges and thus never have the depressing experience of los-

5 Of course, we assume here that the fuse works properly and that there are no other unusual events that intervene.
ing material you have been typing into your computer. Nevertheless, there is a straightforward sense in which you could lose some unsaved material in this manner. After the installation of the surge protector, it is not only true that, were a power surge to occur, the surge protector would prevent your computer from suddenly going off; it is also plausible to think that (now that the surge protector is in place) you cannot lose unsaved material in this manner, even in the absence of a power surge.

Of course, “can” (and “possibility”) and “cannot” (and “impossibility”) are notoriously obscure modalities. But one very plausible way of understanding them is in terms of scenarios of certain sorts. More specifically, in order to support a claim that x can happen (or that x is possible), given circumstances a, b, and c, one must be able to describe a coherent scenario (holding the actual natural laws fixed) in which a, b, and c obtain and x happens. Similarly, if it is true that x cannot happen (or that x is impossible), given circumstances a, b, and c, then there must be no coherent description of such a scenario in which a, b, and c obtain and x happens.

Once the fuse is installed (and supposing it works properly), there is no coherent description of a scenario in which the circuit overheats; thus, the circuit cannot overheat. And once the surge protector is installed (and supposing it works properly), there is no coherent description of a scenario in which you lose unsaved material as a result of a power surge; thus, you cannot lose material in this way. By the same token, once Black is restored to the Frankfurt-style example, there is no coherent description of a scenario in which Jones does other than he actually does; thus, Jones cannot do otherwise.

It is evident, then, that Lamb’s informal reasoning about Frankfurt’s examples is mistaken. Let us now turn to Lamb’s diagnosis of Frankfurt’s error in attempting to argue for a standard interpretation of his examples. Lamb reminds us of Frankfurt’s claim, “But whether [Jones] finally acts on his own or as a result of Black’s intervention, he performs the same action. He has no alternative but to do what Black wants him to do.” Now Lamb says:

I do not know whether Frankfurt intended this passage as an argument, but it can certainly be taken that way, i.e., as the argument that Jones cannot do otherwise because he performs the same action whether or not he acts on his own (521–2).

Lamb proceeds to point out that the general sort of argument suggested here is fallacious. That is, the following general form of argumentation is invalid: either \( p \) is true or \( p \) is not true. If \( p \) is true, then \( x \) occurs. And if \( p \) is not true, then \( x \) occurs. Therefore, \( x \) must occur. This kind of argument is invalid—some sort of modal fallacy—because it may be the case (for example) that even though \( p \) is true, \( x \) might not occur.\(^7\)

We agree that the form of reasoning identified by Lamb is invalid. But we deny that Frankfurt employs it or is in any way committed to it. When Frankfurt says, “But whether [Jones] finally acts on his own or as a result of Black’s intervention, he performs the same action,” he follows this sentence with what we take to be an explication (or elaboration) of it: “He has no alternative but to do what Black wants him to do.” Thus, Frankfurt’s first claim need not be taken as pointing to the obviously fallacious sort of modal argument identified by Lamb. It is much more natural to construe Frankfurt as claiming that the structure of the example is such that there is no alternative accessible to Jones in which he does otherwise (“He has no alternative but to do what Black wants him to do”). And if this is so, then it follows straightforwardly that Jones cannot do otherwise. Of course, this inference is based on the skeletal analysis of “can” (“cannot”) and “possibility” (“impossibility”) suggested above in our discussion of the mistake in Lamb’s informal reasoning about the Frankfurt-style examples. If the structure of the example is such that there is no alternative accessible to Jones in which he does otherwise, then surely there is no coherent description of a scenario in which Black is present (and disposed to act as he actually is disposed to act) and Jones does other than he actually does. Thus, Jones cannot do otherwise.

We believe that our interpretation of Frankfurt is fairer and more natural than Lamb’s. But there is still a bit of a gap in the argumentation. On what basis can Frankfurt claim that the structure of the example renders it true that there is no alternative accessible to Jones in which he does otherwise? Lamb suggests a way in which Frankfurt might seek to establish his claim:

How could Jones have possibly have refrained from what he did? For if he had refrained, the triggering event would have already occurred.

\(^7\) In the specific form of the fallacious pattern of reasoning considered by Lamb, the second conditional premise is more appropriately construed as, “If \( p \) is not true, then \( x \) must occur.” But clearly this does not make any difference to the validity of the argument, since the first conditional premise remains too weak.
And if the triggering event had already occurred, Black would have intervened and forced Jones to act, in which case Jones would not have been able to refrain. Hence, if Jones had refrained, he would have been unable to refrain. But this is absurd. Thus, Jones could not have acted otherwise than he did (523).

Lamb finds this form of reasoning problematic (523–4). He asserts that Jones could have refrained from acting even though it is true that, if Jones had refrained, he would have been unable to refrain. But surely this is an artificial and baroque way of construing Frankfurt’s basis for his view. Rather than employ the unintuitive conditional, ‘If he had refrained, the triggering event would have already occurred’, presumably Frankfurt would employ the conditional, ‘If he were about to refrain (in the absence of intervention by an external agent or factor), the triggering event would have already occurred’. What follows is simply that, if Jones were about to refrain (but for interventions), he would be rendered unable to refrain. And this seems straightforwardly to imply that Jones cannot refrain from doing what he actually does.

We have argued that a certain natural way of understanding the modalities “can,” “possibility,” “cannot,” and “impossibility” helps to expose precisely why Lamb is incorrect in saying that the relevant agents can do otherwise in the Frankfurt examples. A sign that he is incorrect is that this view would commit him to obviously unacceptable results in a class of cases involving the installation of safety devices. Further, this way of understanding the modalities points to a way of interpreting Frankfurt’s own reasoning on behalf of the standard view of these examples (according to which the relevant agent cannot do otherwise); on this interpretation, Frankfurt—or the proponent of the view that his examples call PAP into question—is not guilty of an obvious modal confusion or fallacy.

JOHN MARTIN FISCHER AND PAUL HOFFMAN
University of California/Riverside