
FRANKFURT-TYPE
EXAMPLES AND
SEMI-COMPATIBILISM

JOHN MARTIN FISCHER

It is a basic and pervasive assumption that in order to be morally responsible for one's behavior, one must have had (at some relevant point along the path to the behavior) alternative possibilities of a certain sort. This basic idea is encapsulated in the "Principle of Alternative Possibilities," the various versions of which require that moral responsibility be associated with the presence of alternative possibilities.¹ Now there are powerful reasons to think that causal determinism would rule out alternative possibilities.² So it has appeared to many philosophers that causal determinism is incompatible with moral responsibility.

There are, however, various ways of challenging the Principle of Alternative Possibilities. One way employs a thought-experiment with a distinctive structure; such thought-experiments are frequently called "Frankfurt-type examples," because of Harry Frankfurt's seminal presentation of them (1969). The examples contain a fail-safe mechanism that does not actually play any role in the relevant agent's deliberations, choices, and behavior, but whose presence ensures that the agent deliberates, chooses, and behaves just as he actually does.

1. FRANKFURT-TYPE EXAMPLES

The first “Frankfurt-type case” was given by John Locke in *An Essay Concerning Human Understanding*. Locke’s example is a case in which “a man be carried whilst fast asleep into a room where is a person he longs to see and speak with, and be there locked fast in, beyond his power to get out; he awakes and is glad to find himself in so desirable company, which he stays willingly in . . .”³ In Locke’s example, the man stays in the room voluntarily and it seems that he does so “freely” (although Locke himself would use the term *voluntarily* rather than *freely*) and can be morally responsible for doing so, although unbeknownst to him he could not have left the room. Of course, the man *does* have various alternative possibilities (apart from special assumptions): he can choose to leave the room and try to leave the room, and so forth.

Frankfurt can be seen to be entering the debate at this point. Frankfurt seeks to construct examples in which even *these* sorts of alternative possibilities have been eliminated. To do this, Frankfurt employs the apparatus of a “counterfactual intervener”⁴ who can monitor the brain and intervene in it, should the agent be about to choose to do otherwise. In order to flesh out these examples—although Frankfurt did not explicitly do this—it is useful to posit a “prior sign” that can be read by the counterfactual intervener and guide him in his activity. (This was David Blumenfeld’s innovation: Blumenfeld, 1971.) If the sign indicates that the agent is about to choose to do what the counterfactual intervener wants him to choose, the intervener does not intervene. If, contrary to fact, the agent were about to choose differently, the prior sign would inform the counterfactual intervener (and he would intervene).

Here is a particular version of a “Frankfurt-type case.” Suppose Jones is in a voting booth deliberating about whether to vote for Gore or Bush. After reflection, he chooses to vote for Gore and does vote for Gore by marking his ballot in the normal way. Unbeknownst to him, Black, a liberal neurosurgeon working with the Democratic Party, has implanted a device in Jones’s brain which monitors Jones’s brain activities. If he is about to choose to vote Democratic, the device simply continues monitoring and does not intervene in the process in any way. If, however, Jones is about to choose to vote, say, Republican, the device triggers an intervention that involves electronic stimulation of the brain sufficient to produce a choice to vote for the Democrat (and a subsequent Democratic vote).

How can the device tell whether Jones is about to choose to vote Republican or Democratic? This is where the “prior sign” comes in. If Jones is about to choose at T₂ to vote for Gore at T₃, he shows some involuntary sign—say a neurological pattern in his brain—at T₁. Detecting this, Black’s device does not intervene. But if Jones is about to choose at T₂ to vote for Bush at T₃, he shows an involuntary sign—a different neurological pattern in his brain—at T₁. This brain pattern

would trigger Black's device to intervene and cause Jones to choose at T₂ to vote for Gore, and to vote for Gore at T₃.

In that the device plays no role in Jones's deliberations and act of voting, it seems to me that Jones acts freely and is morally responsible for voting for Gore. And given the presence of Black's device, it is plausible to think that Jones does not have alternative possibilities with regard to his choice and action. Thus, the Frankfurt-type examples seem to be counterexamples to the Principle of Alternative Possibilities.

2. THE "DIVIDE AND CONQUER" STRATEGY OF RESPONSE TO THE FRANKFURT-TYPE EXAMPLES

Peter van Inwagen (1978; 1983) has developed what might be called the "divide and conquer" strategy of response to the Frankfurt-type examples. His basic point is that the proponent of the Frankfurt-type cases (as counterexamples to the Principle of Alternative Possibilities) is not sufficiently precise in specifying what the relevant agent is morally responsible for. We typically hold individuals morally responsible for various items, including actions, omissions, and consequences (envisaged either as "particulars" or more coarsely individuated "universals"). Van Inwagen's contention is that (in the Frankfurt-type cases and elsewhere) there is no one item of which it is true both that there is no alternative to it and that the agent is morally responsible for it. There are some items for which the agent may well be responsible, but these are items to which there are genuinely available alternative possibilities. And there are some items to which there are no such alternative possibilities; for these items the agent is not morally responsible, according to van Inwagen. Van Inwagen's diagnosis of the confusion of the proponent of the Frankfurt-type cases is that he is (perhaps implicitly) thinking of one sort of item when he is focusing on moral responsibility, and another when he is focusing on alternative possibilities.

Van Inwagen essentially distinguishes four principles. The "Principle of Alternate Possibilities" (strictly speaking) deals with actions which have been performed:

- (PAP) A person is morally responsible for what he has done only if he could have done otherwise.

The “Principle of Possible Action” pertains to actions which have not been performed (failures to act or omissions):

- (PPA) A person is morally responsible for failing to perform a given act only if he could have performed that act.

Two additional principles deal with our moral responsibility for consequences of what we do (or omit). The first “Principle of Possible Prevention” applies to consequences considered as event-particulars:

- (PPP1) A person is morally responsible for a certain event-particular only if he could have prevented it.

The second such principle applies to consequences considered as event-universals, which are individuated more broadly than event-particulars:

- (PPP2) A person is morally responsible for a certain state of affairs only if (that state of affairs obtains and) he could have prevented it from obtaining.

Van Inwagen contends that (PPA), (PPP1), and (PPP2) cannot be refuted by Frankfurt-type cases. Further, he claims that if no one is morally responsible for having failed to perform any act, and no one is morally responsible for any event-particular or event-universal, then no one is morally responsible for *anything* (including actions) (1983: 181).

It will be useful to consider the arguments for and against each principle. Let us begin with the principle pertaining to omissions, (PPA). Van Inwagen’s defense of the principle can be understood as follows. Suppose you are in your apartment looking out the window, and you see someone being mugged. You consider calling the police, but you just do not want any involvement, so you refrain from calling the police. Unbeknownst to you, the telephone wire has been cut by the criminal, and so even if you had tried to reach the police, you would have been unsuccessful. Van Inwagen’s intuition is that you are not morally responsible for failing to call the police (in the sense of failing to successfully make contact with the police). You may be morally responsible for failing to *try* to reach the police, for failing to dial, and so forth; but you are not morally responsible for failing successfully to reach the police. Further, it is van Inwagen’s crucial contention that this intuition is explained by the fact that you *could not have reached the police*, and thus that in general moral responsibility for failing to do X requires the ability to do X. Van Inwagen’s strategy involves pointing to a number of cases of omissions in which the agent could not have done the act in question and in which it appears that we would base our exculpation of the agent on this inability.

At one point I accepted this sort of defense of (PPA), but because I reject (PAP), I defended an asymmetry between actions and omissions with respect to the requirement of alternative possibilities for moral responsibility. That is, I defended the thesis that whereas moral responsibility for performing act X does not require the ability to refrain from X-ing, moral responsibility for failing to perform X does require the ability to do X (Fischer 1985/86). But a number of insightful critiques convinced me that in fact (PPA) is just as problematic as (PAP) (Haji 1992; Clarke 1994; Zimmerman 1994; Frankfurt 1994; and Glannon 1995). The basic insight behind the critique of (PPA) is that there are omissions cases that are structurally similar to the Frankfurt-type action cases. Thus, van Inwagen's intuitions may well be valid for the cases (such as the case of the mugging and cut telephone wire described above) to which he points, but this is only a proper subset of the relevant cases.

So, for example, consider an "omissions version" of the Frankfurt-type case with which we began (in which Jones votes for Gore). Suppose everything is the same except that Jones actually refrains from voting at all. (He is disgusted by all the available candidates.) But suppose that Black is ready to directly stimulate Jones's brain, should Jones show a prior sign indicating that he is about to choose to vote (and to vote). It seems to me that Jones can in this case be morally responsible for refraining from voting, although he could not have voted. I do not see any relevant difference between this sort of case and the "action-version" of the example (see Fischer and Ravizza 1998).

Van Inwagen appears to concede that Frankfurt has in fact provided counterexamples to (PAP). Discussing an action version of the Frankfurt-type examples, van Inwagen writes, "It seems we must conclude that we have a genuine case in which an agent is morally responsible for having shot a certain man even though he could not have done otherwise than shoot that man. This case shows that the Principle of Alternate Possibilities is probably false" (1983: 164), I think that van Inwagen is correct here. But I have always been puzzled as to how he can say this, given that he has apparently argued that (PPA), (PPP₁), and (PPP₂), which he accepts, entail that moral responsibility for *anything* requires alternative possibilities.

Let us now turn to the principle as it applies to consequences. It will perhaps be illuminating to begin with the principle pertaining to consequence-universals, (PPP₂):

- (PPP₂) A person is morally responsible for a certain state of affairs only if (that state of affairs obtains and) he could have prevented it from obtaining.

Again, van Inwagen (*ibid.*: 164) develops a case in which the agent could not have prevented the state of affairs from obtaining and it appears that precisely this fact

entails that he is not morally responsible for it. So suppose that Ryder has been kidnapped and involuntarily placed on a horse, Dobbin. Ryder comes to a fork in the road at which he can guide Dobbin to the right or left, but he cannot (at any point) cause Dobbin to stop before he gets to the end of the road. Further, unbeknownst to Ryder, both forks end up in Rome. Van Inwagen's intuition is that Ryder may well be morally responsible for causing Dobbin to take the left rather than the right fork, but he cannot fairly be held morally responsible for the state of affairs that Dobbin ends up in Rome (one way or another). And van Inwagen suggests that precisely the fact that Ryder could not have prevented this state of affairs from obtaining that makes it the case that he is not morally responsible for it.

But, again, as with omissions, it seems to me that Van Inwagen is focusing on a proper subset of cases. I believe that there are other cases in which it is plausible to say that the relevant agent is morally responsible for the obtaining of a state of affairs which he cannot prevent from obtaining. Consider, for example, "Assassin." Sam tells his friend, Jack, of his plan to murder the mayor. Jack also wants the mayor dead, so he has secretly implanted a device in Sam's brain which allows him to monitor all of Sam's brain activity and to intervene in it, if he desires. The device can be employed by Jack to ensure (via direct electronic stimulation of the brain) that Sam decides to kill the mayor and that he acts on this decision. Suppose, further, that Sam methodically and freely carries out his plan to kill the mayor. Jack thus plays absolutely no role in Sam's decision and action; Sam acts exactly as he would have acted had no device been implanted in his brain. It seems to me that in this case Sam is morally responsible not only for his act of shooting the mayor, but for the state of affairs universal, *that the mayor is shot*. And yet (given Jack's set-up) Sam cannot prevent the obtaining of this state of affairs (See Fischer and Ravizza 1998: 59.)

Consider, also, "Missile." Here an evil woman, Elizabeth, has obtained a missile and missile launcher, and she has decided (for her own rather perverse reasons) to launch the missile toward Washington, D.C. Suppose that Elizabeth's situation is like that of Sam; she has not been manipulated, brainwashed, and so forth. Further, imagine that she has had exactly the same sort of device implanted in her brain as had been put into Sam's and that there is a "counterfactual intervener" associated with her who would ensure that Elizabeth would launch the missile, if Elizabeth were to show any sign of wavering. Suppose also that, once the missile is launched toward the city, Elizabeth cannot prevent it from hitting Washington, D.C. When Elizabeth freely launches the missile toward Washington, D.C., it seems to me that she is morally responsible for the occurrence of the consequence-universal, *that Washington, D.C. is bombed*; and yet she could not prevent this state of affairs from obtaining (one way or another).

My contention, then, is that there are indeed cases in which an agent can legitimately be held morally responsible for bringing about a consequence-

universal, even though she could not have prevented this universal from obtaining (one way or another). Van Inwagen's intuition to the contrary here, as in the context of omissions, is based on attending to a proper subset of the relevant cases and inappropriately generalizing from this subset. Further, I (and my co-author) have developed (elsewhere) a principled way of distinguishing among the cases.⁵

Now let us turn to the Principle of Alternative Possibilities as it applies to consequence-particulars:

(PPP₁) A person is morally responsible for a certain event-particular only if he could have prevented it.

Van Inwagen's argument in defense of (PPP₁) proceeds as follows (1983: 167–70). He begins by accepting a criterion of event-individuation according to which the actual causal antecedents of a particular event are *essential* to it. Now Van Inwagen points out that in the alternative sequence in a Frankfurt-type case a different causal sequence (involving a different prior sign) from the actual sequence occurs; thus, according to van Inwagen, there is a different event-particular in the alternative sequence from the event-particular in the actual sequence. Thus, he contends that Frankfurt has not impugned (PPP₁) by providing cases in which the agent is morally responsible and yet the same event occurs in the actual and alternative sequences.

Perhaps van Inwagen's strategy can be understood as follows. The proponent of (PPP₁) essentially believes that moral responsibility requires that the agent have access to an alternative possible world in which a different event-particular results from his behavior. One way to show the falsity of this belief would be to display cases in which the agent is morally responsible and yet the relevant event-particulars are the same in the actual and alternative scenarios. But—given the fine-grained approach to event-individuation—Frankfurt has not succeeded in displaying such a case.

Some philosophers have questioned van Inwagen's "essentialist" principle of event-individuation, and they have thus contended that Frankfurt has in fact succeeded in presenting cases of the requisite sort (for example, Carter 1979). I do not know how exactly to resolve the dispute about event-individuation, but I also do not think that one's views about the Frankfurt-type cases should depend on this sort of issue. Against van Inwagen I would argue that Frankfurt's strategy for impugning (PAP) need not rest on the project of presenting cases in which the event-particulars are the same in the actual and alternative scenarios, because I believe that the proponent of (PPP₁) should hold that moral responsibility requires that the agent have access to an alternative possible world in which a different event-particular results from his *voluntary* behavior. That is, if one believes that one's moral responsibility is grounded in the sort of control that involves genuine

alternative possibilities, it seems to me that those alternative possibilities must contain voluntary behavior; how can adding a scenario in which the agent does not voluntarily bring about a different event-particular make it the case that the agent is morally responsible in the actual sequence? (Fischer 1994: 131–59; see also, Kane 1985: 60 and 1996a: 107–15.) To suppose that this is possible would be to believe in alchemy! Surely, in the Frankfurt-type cases, an agent is, intuitively speaking, morally responsible for bringing about an event-particular, and yet he does not have access to an alternative scenario in which he *voluntarily* brings about a different event-particular. So the Frankfurt-type cases are not best construed as cases of access to a different world with the same event-particular, but of *lack of access* to the *relevant* sort of alternative world. If the Frankfurt-type cases are construed in this way—as showing lack of access to the relevant alternative possibilities—then the issue about individuation of event-particulars becomes irrelevant.

To drive the point home, consider a variant on the Frankfurt-type case of Jones and Black. Here everything is as in the original Frankfurt-type example, except that if Black detects that Jones is about to choose to vote for Bush, Black will use his machine to destroy Jones's brain and thus kill him instantly. Here, again, Jones freely chooses to vote for Gore and does vote for Gore (in the normal way). He behaves just as he would have behaved, had there been no device implanted in his brain. And yet he could not have brought about a different event-particular. In this case it is clear that Jones's lack of ability to bring about a different event-particular does not come from his (sole) access to another possible world in which there is the same event-particular as in the actual world; rather, it comes from his *lack of access* to another possible world of *any* sort (and thus of the *relevant* sort). If I may immodestly dub this sort of case a "Fischer-variant" on the Frankfurt-type cases, I would claim that the Fischer-variants refute (PPP₁) just as effectively as the original Frankfurt-type cases. The distinctive potency of the Frankfurt-type cases consists in showing that there can be moral responsibility even in cases in which the agent *lacks access* to the appropriate alternative scenarios, and thus van Inwagen's ingenious reliance on a fine-grained method of act-individuation to defend (PPP₁) is in the end misguided and irrelevant.

I conclude, then, that none of van Inwagen's arguments in defense of (PPA), (PPP₁), or (PPP₂) is compelling. Indeed, I believe that Frankfurt-type cases provide powerful reasons to reject these principles, along with (PAP).

When considering (PAP), someone might object in a manner similar to the way in which Van Inwagen objected to Frankfurt's critique of (PPP₁). The original case of Jones is supposed to be one in which Jones is morally responsible for his choice and his act of voting for Gore, although he lacks alternative possibilities. At this point it may be objected that, despite the initial appearance, Jones *does* have at least *some* alternative possibility. Although Jones cannot choose or vote

differently, he can still exhibit a different neurological pattern in his brain N^* (from the one he actually exhibits, N). I have called such an alternative possibility a “flicker of freedom” (Fischer 1994). The flicker theorist contends that our moral responsibility always can be traced back to some suitably placed flicker of freedom; our responsibility is grounded in and derives from such alternative possibilities.

It seems that one can always find a flicker of freedom in the Frankfurt-type cases insofar as they are developed as “prior-sign” cases. That is, the agent will always at least have the power to exhibit an alternative sign. But I contend that the mere involuntary display of some sign—such as a neurological pattern in the brain, a blush, or a furrowed brow—is too thin a reed on which to rest moral responsibility. The power involuntarily to exhibit a different sign seems to me to be insufficiently robust to ground our attributions of moral responsibility.

Note that in the alternative sequence (in which Jones shows neurological pattern N^* , which is indicative of an impending decision to vote for Bush), the sign is entirely involuntary and the subsequent decision and vote are produced electronically. Thus, in the alternative sequence, Jones cannot be said to be choosing and acting freely and similarly cannot be thought to be morally responsible for his choice and action. If my point in connection with van Inwagen’s “act-individuation” defense of (PPP₁) is correct, then this sort of alternative possibility cannot ground ascriptions of moral responsibility. It is insufficiently robust: it lacks “voluntary oomph.”

To help to see this, imagine, just for a moment, that there are absolutely no alternative possibilities, even the flimsy and exiguous flickers of freedom we have recently been entertaining. An alternative-possibilities control theorist would say that under such circumstances the relevant agent cannot be morally responsible for his choice and action. Now add the flickers of freedom we have been considering—the power to exhibit a different neurological pattern, N^* . I find it very hard to see how adding this power can transform a situation in which there is no moral responsibility into one in which there is moral responsibility. How can adding a pathway along which Jones does *not* freely vote for Gore and is *not* morally responsible for voting for Gore make it the case that Jones actually *is* morally responsible for voting for Gore? This, again, is the “problem of alchemy.” (Fischer 1994: 141)

I believe that this problem of lack of robustness—lack of voluntary oomph—plagues various versions of the flicker of freedom strategy or response to the Frankfurt-type examples. For example, suppose one follows Margery Bedford Naylor (1984) in arguing that what one is “really” morally responsible for is (say) acting “on one’s own” (and not as a result of coercion, manipulation, and so forth). Now if this is so one could say that the agent does indeed have an alternative possibility—the option of not acting on one’s own in this sense. But I would contend that this sort of alternative possibility is a mere flicker of freedom and insufficiently robust to ground attributions of moral responsibility.

My reason is that in the alternative sequence of a Frankfurt-type case the agent would not be voluntarily choosing not to perform the action on her own. That is, it is true that (in the alternative sequence of a Frankfurt-type case) the agent would not be choosing and acting on her own, but these features of the sequence would not be voluntarily adopted by him—they would be entirely fortuitous, from the point of view of his deliberations. It would then seem to me that the sort of alternative possibility identified by Naylor lacks voluntary oomph.

3. A DILEMMA FOR THE PROPONENT OF FRANKFURT-TYPE EXAMPLES

An important challenge to the position I have sketched (against the flicker theorist) has been presented by such philosophers as David Widerker (1995 a and b), Robert Kane (1985: 51; 1996a: 142–45)⁶, Carl Ginet (1996), and Keith Wyma (1997). I will boil down the various versions of the argument into the following. It begins with a dilemma: either the proponent of the Frankfurt-type examples is presupposing the truth of causal determinism or indeterminism.

Let us start with the presupposition that causal determinism obtains. Now it appears as though the relevant agent, Jones, in the previous example, cannot choose or do otherwise (cannot choose at T₂ to vote for Bush or vote for Bush at T₃) because the “counterfactual intervener,” the liberal neurosurgeon Black, can know, given the prior sign exhibited by Jones at T₁, that Jones will indeed choose to vote for Gore at T₂. If Jones were to choose at T₂ to vote for Bush, the prior sign would have had to be different; thus, Jones cannot at T₂ choose to vote for Bush at T₃. But the problem is that the contention that Jones is morally responsible for choosing to vote for Gore and actually voting for Gore is put in doubt, given the assumption of causal determinism. That is, if causal determinism is explicitly presupposed, it does not seem that someone could say that Jones is obviously morally responsible for his actual choice and action, in a context in which the relationship between causal determinism and moral responsibility are at issue. To do so would appear to beg the question against the incompatibilist.

Now suppose that indeterminism (of a certain relevant sort) obtains. Under this supposition it would not be dialectically inappropriate to claim that Jones is morally responsible for his actual choice at T₂ to vote for Gore and his vote for Gore at T₃. But now the contention that Jones cannot choose at T₂ to vote for Bush at T₃ is called into question, because there is no deterministic relationship between the prior sign exhibited by Jones at T₁ and Jones’s subsequent choice at

T₂. So, if we consider the time just prior to T₂, everything about the past can be just as it is consistently with Jones's choosing at T₂ to vote for Bush at T₃. Someone might think that if it takes some time for Jones to make the choice, Black can intervene to prevent the completion of the choice; but then Jones will still have the possibility of "beginning to make the choice," which is surely more robust than a mere flicker of freedom (say an involuntary twitch, blush, or neurological pattern). After all, beginning to make a choice is a voluntary undertaking (even if it is truncated through no fault of one's own)—it presumably has sufficient voluntary oomph to ground ascriptions of moral responsibility.

The proponents of the Frankfurt-type examples contend that they are non-question-begging cases in which an agent is morally responsible for her choice and action and yet the agent has no sufficiently robust alternative possibilities. But the counter-argument of Widerker, Kane, Ginet, and Wyma appears to show that the examples in question are neither uncontroversial cases in which the agent is morally responsible for his choice and subsequent behavior nor cases in which the agent lacks the alternative possibilities. This clearly important argument has been influential. Indeed, in a recent article Ted A. Warfield (1996: 221) claims that the rejection of the Frankfurt-type examples (as cases in which an agent is morally responsible yet lacks alternative possibilities) is "increasingly common."

4. A REPLY ON BEHALF OF THE PROPONENT OF THE FRANKFURT-TYPE EXAMPLES

Despite this rising chorus I still remain convinced that the Frankfurt-type cases help to establish that moral responsibility does not require alternative possibilities.

4.1. The Assumption of Causal Determinism

Begin with the first horn of the dilemma: the assumption that causal determinism obtains. I agree that one cannot now simply and precipitously conclude, from consideration of the examples, that the agent is morally responsible for his choice and behavior. But in any case this is not the way I would have proceeded; I never have envisaged a simple "one-step" argument to the conclusion that (say) Jones is morally responsible for his choice and action. Rather, I employ the Frankfurt-

type examples as the first (but obviously important) step of a slightly more complex argument to the conclusion that Jones is morally responsible for his choice and action (despite lacking alternative possibilities).

The argument goes as follows. First, one carefully considers the Frankfurt-type cases. Upon reflection, I believe that one should conclude that in these cases the lack of alternative possibilities does not in itself ground a claim that the agent is not morally responsible for his choice and action. In other words, I think that the examples make highly plausible the preliminary conclusion that *if* Jones is not morally responsible for his choice and action, this is *not* simply because he lacks alternative possibilities. After all, everything that has any causal (or any other kind of) influence on Jones would be exactly the same, if we “subtracted” Black entirely from the scene. And Jones’s moral responsibility would seem to be supervenient on what has an influence or impact on him in some way.

So the relevant (preliminary) conclusion is, if Jones is not morally responsible for his choice and action, the reason is not simply that he lacks alternative possibilities. And it does *not* appear to beg the question to come to this conclusion, even if causal determinism obtains. The first step is to argue, based on the Frankfurt-type examples, that intuitively it is plausible that alternative possibilities are irrelevant to ascriptions of moral responsibility. One is supposed to see the irrelevance of alternative possibilities simply by reflecting on the examples. I do not know how to *prove* the irrelevance thesis, but I find it extremely plausible intuitively. When Louis Armstrong was asked for the definition of jazz, he allegedly said, “If you have to ask, you ain’t never gonna know.” I am inclined to say the same thing here: if you have to ask *how* the Frankfurt-type cases show the irrelevance of alternative possibilities to moral responsibility, “you ain’t never gonna know.”

The *second* step in the argument consists in asking whether causal determinism *in itself and apart from ruling out alternative possibilities* threatens moral responsibility. I have considered various possible reasons why someone might think that causal determinism does threaten moral responsibility in itself and apart from ruling out alternative possibilities, and I have come to the conclusion that it is not plausible to accept any of these reasons.⁷ It seems to me that this two-stage argument is highly plausible and does *not* beg the question against the incompatibilist, even on the assumption of causal determinism. Thus I believe that the use of the “prior-sign” cases can be defended against the charge of begging the question.

4.2. The Assumption of Indeterminism

Let us now move to the second horn of the dilemma: the assumption of indeterminism. Here I admit that the prior-sign cases will not describe an agent who

lacks alternative possibilities. But I want to sketch three strategies for modifying the Frankfurt-type case to address this difficulty.⁸

4.2.1. *Hunt's Approach*

A Frankfurt-type case which works as the ones sketched previously in this essay is a “prior-sign” case. But recall that the original “Frankfurt-type” case was presented by John Locke in *An Essay Concerning Human Understanding*. It is important to see that there can be *another* sort of Frankfurt-type case, which takes its cue more closely from Locke’s example; I shall refer to such a case, developed by David Hunt (2000), as a “blockage case.” Note that in Locke’s example the door to the room is actually locked *no matter whether the man is inclined to choose to stay in the room or not*. Imagine, then, that although the actual neural processes in one’s brain (one is here supposing that the mind supervenes on the brain) take place indeterministically, *all other neural pathways are blocked*.⁹ This is a way of bringing the locked door, the blockage, into the brain. Just as in the case of the locked door, the pathways are actually blocked; in contrast to the structure of the prior-sign cases, the pathways’ being blocked does not depend on prior features of Jones. This, then, is a different way of solving precisely the problem Frankfurt sought to solve—one that more simply and naturally takes its cue from Locke. And, importantly, it does *not* appear to introduce alternative possibilities.

4.2.2. *Mele and Robb's Approach*

Here is a second way of modifying the Frankfurt-type cases so that they (allegedly) “work” in a causally indeterministic context. Hunt’s strategy involves “blockage” that is insensitive to prior signs. The second strategy, developed by Alfred Mele and David Robb, involves two simultaneously operating sequences, one of which is indeterministic, the other of which is causally deterministic; the indeterministic sequence actually leads to the result in question, but the deterministic sequence (the operation of which is insensitive to prior signs) would have issued in the same sort of result, if the indeterministic sequence had not. Mele and Robb (1998) develop their ingenious example as follows (changing our cast of characters slightly):

At T_1 , Black initiates a certain deterministic process P in Bob’s brain with the intention of thereby causing Bob to decide at T_2 (an hour later, say) to steal Ann’s car. The process, which is screened off from Bob’s consciousness, will deterministically culminate in Bob’s deciding at T_2 to steal Ann’s car unless he decides on his own at T_2 to steal it or is incapable at T_2 of making a decision (because, e.g., he is dead by T_2 .) (Black is unaware that it is open to Bob to decide on his own at T_2 to steal the car; he is confident that P will cause Bob to decide as he wants Bob to decide.) The process is in no way sensitive to any “sign” of what Bob will decide. As it happens, at T_2 Bob decides on his own to

steal the car, on the basis of his own indeterministic deliberation about whether to steal it, and his decision has no deterministic cause. But if he had not just then decided on his own to steal it, *P* would have deterministically issued, at *T*₂, in his deciding to steal it. Rest assured that *P* in no way influences the indeterministic decision-making process that actually issues in Bob's decision. (ibid.: 101–2)

The actual sequence in the Mele/Robb example is indeterministic, and yet the agent could not have done otherwise due to the unfolding of a deterministic causal sequence that preemptively overdetermines the actual decision. And the relevant agent seems to be morally responsible for his decision and behavior.

4.2.3. *Stump's Approach*

The third strategy for modifying the Frankfurt-type cases to accommodate indeterministic contexts is developed by Eleonore Stump (1990, 1995, 1996a, 1999a, which is a response to Goetz 1999). Stump assumes that there is some sort of one-many correlation between a mental act or state and the firings of neurons in the brain:

When I suddenly recognize my daughter's face across a crowded room, that one mental act of recognition, which feels sudden, even instantaneous, to me, is correlated with many neural firings as information from the retina is sent through the optic nerve, relayed through the lateral geniculate nucleus of the thalamus, processed in various parts of the occipital cortex, which take account of figure, motion, orientation in space, and color, and then processed further in cortical association areas. Only when the whole sequence of neural firings is completed, do I have the mental act of recognizing my daughter. Whatever neural firings are correlated with an act of will or intellect, I take it that in this case, as in all others, the correlation between the mental act and the firing of the relevant neurons is a one-many relation. (Stump 1999a: 417)

On Stump's approach, it is crucial that if the firing of the whole neural sequence correlated with a mental act is not completed, the result is not some truncated or incomplete mental act (say, the beginning of a choice or decision). It is no mental act at all:

If the neural sequence correlated with my recognizing my daughter's face across a crowded room is interrupted at the level of the thalamus, say, then I will have no mental act having to do with seeing her. I won't for example, think to myself, "For a moment there, I thought I saw my daughter, but now I'm not sure." I won't have a sensation of almost but not quite seeing her. I won't have a premonition that I was about to see her, and then I mysteriously just don't see her. I will simply have no mental act regarding recognition of her at all. (ibid.: 417–18)

Let us suppose now that a mental event is identical to a series of neural firings.¹⁰ A particular mental event, say, a choice, can be assumed to result from

an indeterministic process. Further, a counterfactual intervener can be associated with the agent who could notice (in an alternative scenario) that a different neural sequence was beginning, and could then interrupt it before it can be completed. If Black, the counterfactually intervening liberal neurosurgeon, did interrupt a neural sequence that was beginning (and which is such that, if it were completed, it would constitute, or correlate with, a decision to vote for Bush), Jones would *not* (according to Stump) have engaged in the mental act of *beginning to make a decision*. Jones would have *no* mental act, just as Stump would not have begun to recognize her daughter, if the sequence of neural firings beginning in her retina had been terminated in the thalamus (ibid.: 418).

Thus, in Stump's version of the Frankfurt-type cases, the agent's choice is not causally determined, and it is also true that the agent cannot have chosen (or behaved) differently from how she actually chooses (and behaves). And yet it seems entirely plausible that the agent is morally responsible for her choice and behavior in these cases.

Despite the force and influence of the argument (presented by Widerker, Kane, Ginet, and Wyma) against the contention that in the Frankfurt-type cases the agent is morally responsible although he has no alternative possibilities, there is an attractive strategy of response. Even if causal determinism is true, it does not appear to be question-begging to use the cases as part of a two-stage argument (rather than an argument that simply assumes that the relevant agents are morally responsible in the cases). And if causal determinism is false (in certain ways), it still seems (at least at first blush) to be possible to construct versions of the Frankfurt-type cases in which it is plausible to say that the agent is morally responsible and yet lacks alternative possibilities.

5. ANALYSIS OF THE INDETERMINISTIC CASES

It is contentious, however, whether the indeterministic cases presented by such philosophers as Hunt, Mele and Robb, and Stump really work. Let us start by focusing on Hunt's approach. Recall that Hunt envisages a case in which the neural events resulting in the relevant choice are indeterministic, and yet all *other* neural pathways in the brain are "blocked" (as in Locke's "locked-door" example). The question could now be expressed as follows: Does the agent have access to a scenario in which his neural path makes contact with or "bumps up against" the blockage? If so, it would seem that the alternative possibility in question does after

all exist, because if the neural path “bumps up against” the blockage, then presumably the agent is no longer the author of the subsequent act (and is not morally responsible for it).

But how exactly can the agent (or his neural events) bump up against the blockage? It would seem that access to the blockage would require an intermediate set of neural events, different from the actual neural events, that is—as it were, a “bridge” between the actual neural process and the blockage. (In Locke’s example, the agent would have to walk over to the door and try to open it.) But even these intermediate events are presumed to be blocked in Hunt’s example. So it may seem that Hunt has indeed provided an example of the required sort, that is, one in which the agent is morally responsible and yet does not have *any* alternative possibilities.

But the example is difficult to imagine (and thus properly to evaluate). If causal indeterminism obtains in the actual neural pathway, how exactly can it be the case that the agent does not have access to events consisting in bumping up against any of the barriers (intermediate or terminal)? And if the agent really does not have access to any such “bumping” events, how can it be the case that causal determinism does not actually obtain?

Consider the following somewhat rough analogy. Suppose one is driving on a freeway, with some space (as is safe!) between one’s car and other vehicles. But imagine also that all of the off-ramps to the freeway are entirely bottled up with traffic, right from the beginnings of the off-ramps. The spaces between the cars represents that one’s actual driving on the freeway corresponds to causal indeterminism, and the off-ramps’ being blocked points to the lack of alternative possibilities.

But now someone will ask why, if there is indeed space between the vehicles, the driver cannot at least begin to guide his car toward an off-ramp. And if such possibilities of changing direction exist, then these would seem to be alternatives of the relevant sort, that is, characterized by sufficient voluntary oomph. So the example needs to be changed so that one is driving along on the freeway absolutely “up against” the bumpers of the cars in front and back, but not being pushed or pulled in any way by those cars. Of course, if one were being pushed or pulled along, then this would correspond to actual-sequence causal determination. The idea is that it at least seems possible to be driving in such a manner that one is not being pushed or pulled by the contiguous cars and yet (because of the positions of the cars) one does not have the power to change the direction of the car at all. But here again there seems to be the alternative possibility that involves pressure exerted on the contiguous cars. That is, the “bumping events” seem to be ineradicable features of the analogy, and thus it is hard to see how completely to eradicate the “bumping events” from the brain. (For another sort of reply to the blockage strategy, see Robert Kane 2000a.)

David Hunt has also suggested that the context of God's foreknowledge of future events is relevantly similar to Frankfurt-type examples.¹¹ Let us suppose that God exists within the same time framework as humans do, is essentially omniscient, and can know future contingent truths. Let us further assume that causal indeterminism obtains. (Of course, each of these assumptions is contentious, as is their combination.) I believe that it follows from the conjunction of these assumptions (suitably interpreted) that human agents cannot choose or do otherwise; and yet (given certain assumptions about God) God's knowledge plays absolutely no role in human choices and actions. Just as with the "counterfactual intervener" in a Frankfurt-type case, one could "subtract" God from the situation and everything that has a causal impact on the agent's choices and behavior would be exactly the same. If all the preceding is correct, then the context of God's foreknowledge would seem to be one in which an agent could be held morally responsible for her choice and behavior and yet have no alternative possibilities. Here the problem of the apparent ineradicability of the "bumping" events is eliminated, but of course the package of assumptions necessary to do the trick is controversial.

To recapitulate, it seems to me that both the approaches of Hunt and Mele and Robb are promising, but that they posit something contentious: that the actual sequence can be indeterministic and nevertheless absolutely no alternative possibilities exist (even including bumping events). This problem comes out in Mele and Robb at the point at which they contend that the deterministic process *P* "in no way influences" the indeterministic process *X* that actually issues in the decision, and yet that the agent has absolutely no alternative possibility. How exactly is it possible for *P* to "neutralize" all nonactual neural pathways without issuing in causal determination in the actual pathway? I do not think it is obvious that the critiques are decisive, but on the other hand it is unclear whether we have here plausible Frankfurt-type examples that work in indeterministic contexts. God's foreknowledge (envisaged in a certain way) *may* do the trick. Also, if Stump is correct, then "bumping" events may well be insufficiently robust to ground moral responsibility attributions, because the neural bumping events would be insufficient for a mental event with voluntary oomph.¹²

Recently Derk Pereboom has presented an intriguing version of the Frankfurt-type examples which is promising insofar as it appears to work in an indeterministic context. That is, this sort of indeterministic example involves alternative possibilities that *clearly* lack sufficient robustness to justify attributions of moral responsibility. Here is Pereboom's case:

Joe is considering whether to claim a tax deduction for the substantial local registration fee that he paid when he bought a house. He knows that claiming the deduction is illegal, that he probably won't be caught, and that if he is, he can convincingly plead ignorance. Suppose he has a very powerful but not al-

ways overriding desire to advance his self-interest no matter what the cost to others, and no matter whether advancing his self-interest involves illegal activity. Furthermore, he is a libertarian free agent. But his psychology is such that the only way that in this situation he could choose not to engage in the tax evasion is for moral reasons. His psychology is not, for example, such that he could decide not to evade taxes for no reason or simply on a whim. In fact, it is causally necessary for his deciding not to evade taxes in this situation that a moral reason occur to him with a certain force. A moral reason can occur to him with that force either involuntarily or as a result of his voluntary activity (e.g. by his willing to consider it, or by his seeking out a vivid presentation of such a reason). But a moral reason occurring to him with such force is not causally sufficient for his deciding not to evade taxes. If a moral reason were to occur to him with that force, Joe could, with his libertarian free will, either choose to act on it or act against it (without the intervener's device in place). But to ensure that he decide to evade taxes, a neuroscientist now implants a device which, were it to sense a moral reason occurring with the specified force, would electronically stimulate his brain so that he would decide to evade taxes. In actual fact, no moral reason occurs to him with such force, and he chooses to evade taxes while the device remains idle.¹³

In Pereboom's version of the Frankfurt-type examples, the actual sequence is indeterministic, and the alternative possibility in question—the occurrence to the agent of a moral reason with a sufficient force—does not appear to be sufficiently robust to ground ascriptions of moral responsibility. This is in part because the occurrence of such a reason is not sufficient in itself for Joe to begin to *act* in accordance with it. The possibility of the mere occurrence of a reason (of a certain force) to an individual does not yet ground a claim of the possession of alternative possibilities sufficiently robust to ground moral responsibility. Thus Pereboom's example is promising as a way of both securing actual-sequence indeterminism and alternative possibilities without oomph.

6. A REPLY ON BEHALF OF THE FLICKER THEORIST: ESCAPABILITY OF AUTHORSHIP/RESPONSIBILITY

The critic of the Frankfurt-type examples has (at least) one more card to play. Consider, for example, the following remarks of Michael McKenna (1997):

[W]hat intuitively drives [the proponent of the alternative-possibilities requirement] 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 . . . 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. . . . (ibid.: 73–74)

McKenna elaborates as follows:

The issue . . . here is whether the will . . . places *my* stamp upon the world, and whether *it is up to me* . . . 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 . . . 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? (ibid.: 74–75)

McKenna is claiming that even in the Frankfurt-type cases, the relevant agent has a significant and robust power: the power either to be the author of his action or not, and thus the power to be morally responsible for his action or not. A similar point is made in recent article by Keith Wyma (1997). Wyma begins with an example which suggests that many of us experienced something like a Frankfurt-type example as we were growing up:

When I was four years old and learning to ride a bicycle, I reached a point where my father decided I no longer needed training wheels. But he still worried that I might fall. So on my first attempt “without a net,” he ran alongside as I pedaled. His arms encircled without touching me, his hands resting lightly upon me, but not holding me upright. I rode straight ahead. My father did not push or guide me, but if I had faltered or veered suddenly to the side, he would have tightened his grip, keeping me vertical and on track. After finally braking to a stop, I was jubilant but somewhat hesitant over whether I should be. I wondered, had I really ridden my bike on my own? . . . Was the triumph of riding straight down the street mine or not? (ibid.: 57)

Wyma goes on to argue for an intuition very similar to McKenna's. On Wyma's view, moral responsibility requires a certain kind of “leeway.” And this leeway is specified by what Wyma calls the “Principle of Possibly Passing the Buck” (PPPB):

A person is morally responsible for something she has done, *A*, only if she has failed to do something she could have done, *B*, such that doing *B* would have rendered her morally non-responsible for *A*. (ibid.: 59)

Of course, in a Frankfurt-type case the relevant agent would not be morally responsible in the alternative sequence; Jones would not be morally responsible for

voting for Gore, in the circumstance in which Black's device were triggered. Thus Wyma has apparently identified a significant sort of "leeway," even in the Frankfurt-type examples. At the end of his article, Wyma returns to the analogy with which he started:

I believe the bike riding triumph *was* mine, because even though I could not have fallen or crashed while my father hovered protectively over me, I could still have faltered enough that he would have had to steady me; and because I had leeway to falter but did not do so, the success of riding was truly mine. *PPPB* vindicates a similar kind of leeway as being necessary for ascriptions of moral responsibility. (ibid.: 68)

Additionally, Michael Otsuka has recently defended a principle similar to Wyma's Principle of Possibly Passing the Buck. Otsuka calls his principle the "Principle of Avoidable Blame":

One is blameworthy for performing an act of a given type only if one could instead have behaved in a manner for which one would have been entirely blameless. (Otsuka 1998: 688)¹⁴

Thus, all three defenders of the alternative possibilities control requirement seem to be pointing to the same sort of alternative possibility, they claim is present quite generally, and hence in the Frankfurt-type examples. This is the freedom to "pass the buck" or "escape" or "avoid" moral responsibility. And it seems that this freedom is present in all of the modifications of the Frankfurt-type examples presented here. One might say that these theorists are seeking, perhaps with some success, to fan the flickers of freedom.¹⁵

7. A FURTHER REPLY ON BEHALF OF THE PROPONENT OF THE FRANKFURT-TYPE EXAMPLES

But I believe that problems similar to those of the earlier defenses of the alternative-possibilities control requirement also plague the new approaches. Recall that the problem with saying that the possibility of exhibiting a different prior sign or indicator of future decision (and action) grounds moral responsibility is that the envisaged possibility is too exiguous and flimsy. The displaying of such

a sign would not even be voluntary behavior. How could moral responsibility rest on such a delicate foundation?

Now it might be thought that the possibility of avoiding authorship or the possibility of avoiding moral responsibility would be a more substantial basis for moral responsibility. But I believe there are similar problems here. Note that in the alternative sequence in a Frankfurt-type case the agent would indeed be avoiding, say, moral responsibility, but she would be doing so “accidentally.” The agent would *not* be *voluntarily* avoiding responsibility. The suggestion that avoiding responsibility is a sufficiently robust basis for moral responsibility may derive some of its plausibility from the fact that in a typical context in which we would say that someone has avoided, say, blameworthiness, he would have performed some voluntary action. Typically, the relevant facts about the various paths available to the agent would be accessible to him, and he would voluntarily choose a right action (rather than a morally objectionable one). Here we would say that the agent avoided blameworthiness; but this is a very different sort of context from the Frankfurt-type cases. In the Frankfurt-type cases, the agent does not choose to be morally responsible rather than not—these issues play no role in his deliberations. And in the alternative scenario in a Frankfurt-type case, the agent does not choose to escape responsibility or voluntarily choose anything that implies her escaping responsibility.

To isolate this point out more clearly, note that in the alternative scenario in a Frankfurt-type case the agent does not deliberate about whether or not to embrace moral responsibility. So issues about whether or not to be morally responsible play no explicit role in his deliberations. Further, they play no “implicit” role either. They might play an implicit role in the sort of context discussed above, in which an agent has internalized certain norms on the basis of which he chooses to do what he takes to be the right action. If he successfully avoids blameworthiness here, it is partly attributable to his having internalized norms the relevant community shares. Given these norms, the agent can reasonably expect to escape blame, if he chooses as he does. But in the alternative scenarios in the Frankfurt-type cases issues about moral responsibility obviously do not play an implicit role of this sort.

To the extent that issues pertaining to moral responsibility play neither an explicit nor an implicit role, I shall say that moral responsibility is not “internally related” to the agent’s behavior in the alternative sequence of a Frankfurt-type case. And my point is that it is very plausible that moral responsibility must be so related to the agent’s behavior, in order for the alternative possibility in question to be sufficiently robust to ground ascriptions of moral responsibility.

Of course, I do not accept the alternative-possibilities control model of moral responsibility. But my contention is that, *if* you do buy into this traditional pic-

ture, then you should *also* acknowledge that the alternative possibilities must be *of a certain sort*—they must be sufficiently robust. (I developed this point in my discussion of van Inwagen’s defense of PPP₁ earlier.) This same point has been highlighted by a philosopher with a very different orientation from mine: Robert Kane (1985: 60, 1996a: 107–15), a libertarian who believes that alternative possibilities are required for moral responsibility. Kane emphasizes what he calls the “dual” or “plural” voluntariness (and responsibility) conditions on moral responsibility: the relevant alternative possibilities—that is, alternative possibilities sufficiently robust to ground moral responsibility—must themselves involve voluntary behavior (for which the agent is morally responsible). On Kane’s picture, it is not enough that an agent have *just any sort of alternative possibility*; it must be an alternative in which the agent acts voluntarily and is morally responsible. Similarly, I would contend that the relevant alternative possibilities must contain voluntary, responsible behavior in which moral responsibility is internally related to the agent’s behavior. My suggestion, then, is that the new defenses of the alternative-possibilities control requirement (presented by McKenna, Otsuka, and Wyma) fall prey to the same sort of problem that afflicted earlier such defenses: the alternatives they postulate are not sufficiently robust.

In my early essay, “Responsibility and Control” (1982), I argued that the critic of the Frankfurt-type case mixes up “possibility” and “ability” in a certain way. That is, I pointed out that even if another event (or set of events) occurs in the alternative sequence of a Frankfurt-type case, it does not follow that the agent has the *ability* (in the relevant sense) to bring about this alternative event (or set of events). I believe that the recent defenses of alternative-possibilities control simply reinscribe the same general problem. The lack of “internal relatedness” of moral responsibility to the events in the alternative sequence points to the fact that the agent lacks the relevant sort of ability, even if there exists the possibility of something different happening. So, even if there exists the possibility that the agent not be the author of his action (or avoid moral responsibility), it does not follow that the agent has the ability (in the relevant sense) to avoid authorship (or responsibility). It is a simple point that has played a crucial role in discussions of indeterministic conceptions of control and moral responsibility: the mere possibility of a different event’s occurring does *not* entail that the agent has the *ability* to do otherwise. The point applies equally in the context of the Frankfurt-type examples.

Return to Wyma’s striking claim about his early bike-riding experience, “I believe the bike riding triumph *was* mine, because even though I could not have fallen or crashed while my father hovered protectively over me, I could still have faltered enough that he would have had to steady me; and because I had leeway to falter but did not do so, the success of riding was truly mine” (Wyma 1997: 68). Whereas we could quibble endlessly about details of these sorts of examples,

the *intuitive point* seems clear: it is *not* the possibility of faltering slightly that makes the young Wyma's bike-riding triumph truly his. This has to do *not* with whether he could have faltered slightly, but with how he rode the bike—how he moved the pedals, balanced, and so forth, and by what sort of causal process this all took place.

8. THE “NONSTANDARD RESPONSE” TO THE FRANKFURT-TYPE CASES

What might be called the “standard” critique of the Frankfurt-type cases concedes that the relevant agent does not have available the sort of alternative possibility typically associated with responsibility, but claims that nevertheless one can find a suitable alternative possibility, even in the Frankfurt-type examples. The “nonstandard” response to the examples claims that even in the Frankfurt-type cases the agent has available a robust alternative possibility of precisely the sort normally associated with moral responsibility. So, for example, in the Frankfurt-type example with which we began (in which Jones chooses to vote for Gore and votes for Gore on his own but Black is poised to intervene should Jones show any sign of choosing to vote for Bush), the nonstandard response claims that Jones *does* have the ability (in the relevant sense) to vote for Bush (Lamb 1993; Campbell 1997; for replies, see Fischer and Hoffman 1994 and McKenna 1998a).

The basic claim of the nonstandard response is that insofar as Black's device does not play any *actual* role in Jones's deliberations or actions, it should be “subtracted” when one is considering whether Jones has a genuine ability to choose and do otherwise (to choose to vote for Bush and to vote for Bush). Put in terms of possible-worlds semantics, the compatibilist in general insists that the alternative possible worlds that establish that the agent has a certain power or ability can *differ* in various ways from the actual world. So, if a given agent does X rather than Y in the actual world, but it is intuitively true that the agent could have done Y instead, one looks for a possible world that can *differ* in certain ways from the actual world prior to the time in question in which the agent does indeed do Y. Given that Black's device does not actually play any role, the nonstandard theorist will say that it can be absent from a possible world that is nevertheless relevant to whether Jones actually has the power to choose to vote for Gore and the power to vote for Gore.

I have difficulty sympathizing with this response. First, note that it would seem to imply that the man who is, unbeknownst to him, locked in the room in Locke's example *can* leave the room. Or, similarly, if one has secretly been chained to one's desk (perhaps while asleep), it follows that one *can* leave one's office, despite the chains, as long as one is unaware of the chains. Also, if one has been struck by paralysis but is not yet aware of the paralysis, the nonstandard approach would seem to imply that the paralysis is no obstacle. But these results just seem very implausible.

I would offer the following diagnosis of the confusion of the nonstandard response. I believe that the nonstandard theorist is conflating general abilities with the sort of ability that corresponds to J. L. Austin's "all-in sense of 'can,'" or "can in the particular circumstances." One may have a general ability without having the latter sort of ability, insofar as one does not have the opportunity to exercise the general ability. Whereas I would certainly concede that the agent in a Frankfurt-type case has the relevant general ability, I would deny that he can under the particular circumstances choose and do otherwise. So, whereas Jones certainly has the general ability to choose to vote Republican and to do so, he cannot under the particular circumstances choose to vote Republican and vote Republican. It is the particularized notion of ability, and not the general one, that is typically associated with moral responsibility; certainly, the particularized notion plays a role in the Principle of Alternative Possibilities.

9. SOME PUTATIVE IMPLICATIONS OF THE FRANKFURT-TYPE CASES

There has (obviously) been considerable ink spilled over the Frankfurt-type examples. Recently some philosophers have explored some previously unnoticed (or insufficiently noticed) aspects of the examples. The Frankfurt-type examples purport to be contexts in which an agent can be morally responsible, even though she lacks alternative possibilities. But consider the following argument, which has been presented and discussed recently by David Widerker (1991a), David Copp (1997), and Ishtiyaque Haji (1993).¹⁶ Suppose someone does something that is intuitively "bad," such as lying just to bolster his reputation. If this act is blameworthy, then it must be wrong. And if it is wrong, it must be the case that the agent should have done something else instead (where this could include simply refraining from doing anything). But "ought implies can," so if the agent

should have done something else instead, then he must have been able to do something else. Thus, if the agent had no alternative possibilities (and thus could *not* have done anything else), then his act of lying cannot be considered blameworthy.

This sort of argument threatens the idea that an agent can be genuinely blameworthy in a context in which he has no alternative possibilities. Insofar as an account of moral responsibility will certainly need to accommodate agents' blameworthiness on some occasions, the argument casts into doubt whether an adequate "actual-sequence" account of moral responsibility can be given.¹⁷

The argument is disturbing, and worthy of more careful attention than I can give it here. Instead, I briefly suggest various ways of responding to the argument. One might deny the maxim that "ought implies can." Various philosophers have rejected it on grounds independent of considerations pertinent to the Frankfurt-type cases. Typically these philosophers have been motivated to give it up in light of reflection on the logic of moral dilemmas.¹⁸ Another approach is suggested by Haji, who rejects the contention that if an act is blameworthy then it is "objectively wrong." That is, the argument presupposes that blameworthiness is connected to objective wrongness, whereas Haji believes that blameworthiness is linked only with subjective wrongness. So, on Haji's view, if an agent is blameworthy for performing an action, it need not be the case that the action was wrong, only (roughly) that the agent believed it to be wrong (and nevertheless did it).¹⁹

My own inclination here is to reject the "ought implies can" maxim, according to which, if an agent ought to do X, then he can do X. But why exactly should one accept this maxim? That is, what justification could be offered for it? It is most natural, I think, to say that it is valid because if it were not, then there could be cases in which an agent ought to do X but in fact cannot do X (and never could do X). Thus, given the connection between its being the case that an agent ought to do X and the agent's being blameworthy for not doing X, there could be cases in which an agent is blameworthy for not X-ing and yet he cannot X. And this seems unfair.

But I argued earlier that some Frankfurt-type omissions cases are relevantly similar to Frankfurt-type cases with respect to actions. That is, there are cases in which an agent is morally responsible for not X-ing although he cannot in fact X. Some of these are cases in which an agent is blameworthy for not X-ing and yet he cannot X. In fact, I believe that anyone who accepts the Frankfurt-type action cases must accept that there are such omissions examples. Thus, precisely the basic intuitions elicited by the Frankfurt-type cases show that the most natural justification of the "ought implies can" maxim is faulty. It is therefore not ad hoc for anyone who accepts the standard interpretation of the Frankfurt-type cases to reject the "ought-implies can" maxim.²⁰

10. AN ACTUAL-SEQUENCE APPROACH TO MORAL RESPONSIBILITY

In my view, the Frankfurt-type cases provide very strong reasons to think that moral responsibility does not require alternative possibilities. Of course, they fall short of providing *decisive* reason to abandon the Principle of Alternative Possibilities. But they should make a reasonable person abandon an endless attempt seeking to identify some sort of alternative possibility and instead set about identifying what about the actual sequence of events leading to an action (or omission or consequence) grounds ascriptions of moral responsibility.

The lesson of the Frankfurt-type cases is that in assessing moral responsibility we should focus on the properties of the actual sequence of events leading to the behavior in question. Of course it does not follow straightforwardly that causal determinism is indeed compatible with moral responsibility. It is still possible to argue for “actual-sequence incompatibilism”—the view that causal determinism rules out moral responsibility quite apart from ruling out alternative possibilities.²¹

My position here is that the argument for the incompatibility of causal determinism and alternative possibilities is *considerably stronger* than the argument that causal determinism rules out moral responsibility *directly* (that is, apart from considerations pertaining to alternative possibilities). I believe that reasonable people, not already committed to a particular position on the free will debate, would find it highly plausible that causal determinism rules out alternative possibilities. As pointed out previously, the argument here proceeds from such plausible principles of common sense as the Principle of the Fixity of the Past and the Principle of the Fixity of the Natural Laws. In contrast, I do not see that any considerations would move a reasonable and fair-minded person not already committed to a particular position on the free will debate, to believe that it is highly plausible that causal determinism rules out alternative possibilities *directly*.²²

Now I do not wish to contend that considerations cannot be invoked which claim that causal determination in the actual sequence rules out moral responsibility directly (and thus not by ruling out alternative possibilities). But I believe that these considerations do not have the broad appeal of those that seem to show that causal determinism rules out alternative possibilities. Further, given that there are strong motivations toward compatibilism between causal determinism and moral responsibility—especially the desire to protect our status as morally responsible agents from esoteric scientific discoveries about the form of the equations that describe the universe—I am inclined to adopt “semi-compatibilism,” the doctrine that causal determinism is compatible with moral responsibility, even if causal determinism were to rule out alternative possibilities.

Of course, there are various ways of specifying and developing an “actual-sequence” approach to moral responsibility (see Fischer 1999a). My approach contends that when one “decodes” the information embedded in the actual sequence in which there is moral responsibility, one will find a certain sort of “control.” Whereas typically it is thought that control must involve alternative possibilities, I believe there are two species of control. “Regulative control” does indeed involve alternative possibilities, but “guidance control” does not; guidance control is of the sort displayed by agents in the actual sequences of Frankfurt-type examples, and, in general, by agents who are morally responsible for their behavior.

My project has been to analyze of guidance control and to show that this sort of control is compatible with moral responsibility. (See Fischer 1994 and Fischer and Ravizza 1998.) In my view, guidance control of one’s behavior has two components: the behavior must issue from one’s own mechanism, and this mechanism must be appropriately responsive to reasons. I have sought to provide accounts of both components, mechanism ownership and reasons-responsiveness, and I have defended the idea that guidance control, so analyzed, is compatible with causal determinism.

NOTES

.....

This chapter builds on—and relies considerably on—previously published work. In particular, I am grateful for permission from the University of Chicago Press to reprint parts of Fischer (1999a), and from Kluwer Academic Publishers to reprint parts of (1999b).

1. Some philosophers prefer “alternate possibilities,” whereas others prefer “alternative possibilities.” Harry Frankfurt offers a (somewhat curmudgeonly) defense of his use of the term “alternate possibilities” in Frankfurt (1999: 372).

2. See, for example, Ginet (1966 87–104) and Ginet (1990); Wiggins (1973: 31–62), van Inwagen (1983), and Fischer (1994).

3. Bk. II, ch. 11, sec. 10.

4. This term was introduced in Fischer (1982).

5. See Fischer and Ravizza (1998: 92–122). For helpful discussions of moral responsibility for consequence-universals, see Heinaman (1986) and Rowe (1989).

6. As far as I know, Kane was the first to articulate this strategy in reply to the Frankfurt examples: Kane (1985: 51).

7. Fischer (1994: 147–54). For further discussion of this issue, see Kane (1996a: 40–43) and Mele (1996: 123–41).

8. For yet another approach, see Fischer (1995), Widerker and Katzoff (1996), Hunt (1996a), and Speak (1999).

9. I borrow this example from Hunt (2000). He develops this—and related—examples further in, “Freedom, Foreknowledge, and Frankfurt” (n.d.).

10. This supposition is just for simplicity's sake; Stump's view is compatible with other stories as to the precise relationship between mental states and brain events.

11. Hunt, working paper.

12. It should be noted that the original proponents of the "indeterminist" strategy of reply to the Frankfurt examples, such as Kane and Widerker, have attempted to respond to the challenges of Fischer, Hunt, Mele and Robb, and Stump. Kane attempts to respond to Hunt and Fischer on blockage cases in Kane (2000a). He attempts to respond to Mele and Robb in "Responsibility, Incompatibilism and Frankfurt-Style Examples," unpublished manuscript to be published in a collection edited by McKenna and Widerker, and to Stump in (Kane 2000b). Widerker attempts to respond to all three strategies in Widerker (2000a).

13. Pereboom, (2001). (Chapter 2 of the book also contains a critical discussion of the Hunt, Mele/Robb, and Stump strategies.)

14. Otsuka qualifies the principle to apply to cases in which it is not the case that everything one is capable of doing at a given point in time is blameworthy because of some previous choice for which one is to blame.

15. Thanks to Dan Speak for this phrase.

16. There is an instructive, extended discussion of these issues in Haji (1998: 42–64 and 151–67).

17. Haji has employed a similar argument to call into question whether morality itself could exist in a world without alternative possibilities: Haji (1998: 42–54).

18. For a thorough discussion, see Sinnott-Armstrong (1998), who believes that the "ought-implies-can" maxim is not an entailment but rather functions as a conversational implicature.

19. There is a critical discussion of this view in Copp (1997).

20. I thank Mark Ravizza for helping me to see this point. Note that the Widerker/Copp argument is a challenge for *any* account of moral responsibility according to which responsibility does not require alternative possibilities, not just to those accounts motivated by Frankfurt-type examples. As I pointed out in the text, consideration of the Frankfurt-type cases provides a powerful way of rejecting the Widerker/Copp argument; it is not clear that one who wishes to eschew such examples has a similarly potent response.

21. Fischer (1982) and Kane (1996a). (Despite his status as an incompatibilist, Kane agrees with me that such a strategy ultimately will not work.)

22. Fischer forthcoming.