



## Philosophical Review

---

Rescuing Frankfurt-Style Cases

Author(s): Alfred R. Mele and David Robb

Reviewed work(s):

Source: *The Philosophical Review*, Vol. 107, No. 1 (Jan., 1998), pp. 97-112

Published by: [Duke University Press](#) on behalf of [Philosophical Review](#)

Stable URL: <http://www.jstor.org/stable/2998316>

Accessed: 26/03/2012 17:38

---

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).



Duke University Press and *Philosophical Review* are collaborating with JSTOR to digitize, preserve and extend access to *The Philosophical Review*.

<http://www.jstor.org>

## Rescuing Frankfurt-Style Cases

Alfred R. Mele and David Robb

Almost thirty years ago, in an attempt to undermine what he termed “the principle of alternate possibilities” (*PAP*),

*PAP*: A person is morally responsible for what he has done only if he could have done otherwise (1969, 829),

Harry Frankfurt offered an ingenious thought-experiment.

*The thought-experiment*. 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. . . . [However] 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. (1969, 835–36)<sup>1</sup>

If, as Frankfurt suggests, Jones could not have done otherwise but is morally responsible for deciding and acting as he did, *PAP* is false.

This objection to *PAP* has played a major role in subsequent literature on moral responsibility and freedom of choice and action. Most contributors to that literature have deemed the objection successful. Compatibilists have used it in criticizing incompatibilism about moral responsibility and about the kind of freedom that such responsibility requires,<sup>2</sup> and some incompatibilists have attempted to accommodate Frankfurt’s moral in refined incompatibilist views.<sup>3</sup> However, several philosophers have argued re-

---

<sup>1</sup>In reproducing this passage, we deleted a subscript after ‘Jones’.

<sup>2</sup>See, for example, Fischer 1994, chap. 7.

<sup>3</sup>In this connection, Widerker (1995b, 113) cites Stump 1990, Stump and Kretzmann 1991, and Zagzebski 1991. He also cites a suggestion made in Fischer 1982. (Fischer is not a libertarian.)

cently that Frankfurt's example and examples like it—"Frankfurt-style" examples—are fundamentally flawed.<sup>4</sup>

Our aim in this paper is to show that a recent and seemingly devastating way of attacking Frankfurt-style examples fails to undermine this general style of attempted counterexample to *PAP*. We will concentrate on detailed objections raised by Robert Kane and by David Widerker (and indicate in notes how our reply to them defuses other recent objections).

### 1. Kane and Widerker on Frankfurt-Style Cases

For the sake of specificity, we will suppose that what Jones decided to do in Frankfurt's case was to steal Ann's car and that Jones successfully executed this decision. Bear in mind that for a traditional incompatibilist about determinism and moral responsibility, if Jones was deterministically caused to decide to steal Ann's car, then he is not morally responsible for deciding to do this. So if Frankfurt's case is to persuade a traditional incompatibilist that Jones is morally responsible for deciding to steal Ann's car, it must be a feature of the example that Jones was not deterministically caused so to decide. We assume accordingly that this is a feature of the example.

Robert Kane argues that in a scenario of this sort, "the controller [that is, Black] cannot know" in advance whether Jones will decide on his own to steal the car (1996, 142). After all, what, if anything, Jones decides on his own to do is supposed to be causally undetermined. Now, either Black intervenes or he does not. If Black intervenes and makes Jones decide to steal the car, then (other things being equal) Jones is not morally responsible for so deciding. And, Kane contends, if Black "does not intervene to predetermine the outcome and the indeterminacy remains in place until the choice [or decision] is made—so that the outcome is [a "self-forming willing" (for example, a causally undetermined decision)]—then the agent . . . is ultimately responsible for it. But then it is also the case that the agent *could have done otherwise*" (142).<sup>5</sup> He could have done otherwise than decide to steal the car because,

---

<sup>4</sup>See Lamb 1993; Widerker 1995a and 1995b; Kane 1996, 142–43, 191–92; Ginet 1996; Wyma 1997.

<sup>5</sup>For a list of various kinds of "self-forming willing," see Kane 1996, 125.

for example, given that Black does not intervene, and given that it is causally undetermined whether he will decide to steal the car and causally undetermined as well whether he will decide *not* to steal the car, it is causally open to Jones to make either decision. Thus, Kane contends, “the simple fact of indeterminacy right up to the moment of choice” or decision undermines Frankfurt’s challenge to incompatibilists (234 n. 23).

In short, Kane confronts Frankfurt with a dilemma. Either (1) Black intervenes and makes Jones decide to steal the car, in which case (other things being equal) Jones is not morally responsible for deciding to steal the car, or (2) Black does *not* intervene, in which case, given that what Jones decides is causally undetermined, Jones could have done otherwise than decide to steal the car. In either case, *PAP* is not threatened (cf. Kane 1985, 51 n. 25; Wyma 1997).

Kane’s argument resembles an argument advanced by David Widerker in a pair of papers (1995a, 1995b).<sup>6</sup> Widerker remarks that “the success of Frankfurt’s case against *PAP* depends crucially upon his ability to convince us of the plausibility of” the following claim:

*IRR*: There may be circumstances in which a person performs some action which although they make it impossible for him to avoid performing that action, they in no way bring it about that he performs it. (1995a, 248; cf. Ginet 1996, 409)

A Frankfurt-style case will succeed in falsifying *PAP* only if it includes circumstances of this kind; but, Widerker argues, no Frankfurt-style case includes such circumstances.

Unlike Kane, Widerker pays special attention to the idea that the Frankfurt-style counterfactual controller (Black) uses a *sign* as a basis for deciding whether to intervene (cf. Wyma, 63–67). If Black is to know when to intervene and when not to do so, it seems that he needs to be receptive to some sign of what Jones is, to use Frankfurt’s expression (835), “going to decide.” So suppose that Black knows that if Jones were to blush at *t1*, that would be a sign that he will do at *t2* what Black wants him to do at *t2*—for example, decide to steal Ann’s car (cf. Widerker 1995a, 249). (Like Kane and Widerker, we regard deciding to *A* as an action—a mental

---

<sup>6</sup>Kane comments on Widerker’s argument in his 1996 (233 n. 22).

action.) Suppose further that Jones does blush at  $t1$ , that Black accordingly does not interfere, and that, at  $t2$ , Jones decides to steal Ann's car (and later executes that decision). Now, if Jones's blushing at  $t1$  deterministically causes his deciding to steal the car, or is in some way associated with something at  $t1$  that deterministically causes this act of deciding, the Frankfurt-style case has gone awry, for reasons we identified earlier (cf. Widerker 1995a, 250; Ginet 1996, 408–9). So suppose that there is no such deterministic causation in the scenario. In that case, Widerker contends, "it is hard to see how Jones's decision is unavoidable" (1995a, 251). Perhaps, given that Jones is blushing at  $t1$ , it is *probable* that he will decide to steal the car (Widerker 1995a, 250). But if what we have here is merely a probability (less than 1), then the following combination is possible: Jones blushes at  $t1$ , Black accordingly does not intervene, and Jones decides at  $t2$  not to steal the car (Widerker 1995a, 250–51).<sup>7</sup> Thus, in at least one intelligible sense of 'could have done otherwise', Jones could have done otherwise at  $t2$  than decide to steal the car, and PAP is not falsified by the case (on this reading of 'could have done otherwise').

In short, Widerker presents Frankfurt with a dilemma. Either (1') the sign Black uses is (associated with) a deterministic cause of Jones's deciding to steal the car, in which case the scenario is not "an *IRR*-situation" or (2') the sign is *not* (associated with) a deterministic cause of Jones's deciding to steal the car, in which case it is possible for Jones to have done otherwise than decide to steal the car (1995a, 251; cf. Wyma, 66–67).

## 2. Reply

In a more ambitious paper than ours, a variety of replies to the foregoing might be explored. One might attack the second horn of Widerker's and Kane's dilemmas. Perhaps even in the case of a causally undetermined decision, it is possible for some hypothetical being (for example, a god) to know whether an agent will decide

---

<sup>7</sup>Widerker seems to suggest (1995a, 251) that Lamb 1993 advances an argument of this kind. We understand Lamb differently, and one of us has replied to Lamb's argument elsewhere (Mele 1995, 101, 141–42). But if Widerker is right about Lamb, our reply to Widerker and Kane is a reply to Lamb as well. (For a detailed reply to Lamb, see Fischer and Hoffman 1994.)

on his own at  $t$  to  $A$ , or whether the agent would so decide at  $t$  if the hypothetical being were not to intervene (cf. Fischer 1995). And perhaps it is possible for some such hypothetical being to be an irresistible intervener. Alternatively, one might attack the first horn of Widerker's dilemma. It may be that cases can be described in which a Frankfurt-style counterfactual controller is sensitive to a perfectly reliable deterministic sign and, even so, some incompatibilists would be willing to grant that the agent is morally responsible for his decision (cf. Fischer 1995). Suppose, for example, that on the basis of painstaking, rational reflection, wealthy Smith judges it best to leave all her money to your favorite cause in her will, and that, given that she so judges, it is causally determined that she will decide to do this; but whether she would judge this best was causally undetermined, and if she had not judged this best, Black would have made her decide to do what she actually decided on her own to do.<sup>8</sup>

We have elected not to pursue these lines of reply. Instead, we offer a Frankfurt-style scenario that goes between the horns of both dilemmas. The scenario is immune to Kane's and Widerker's objections to Frankfurt-style cases.

Our scenario features an agent, Bob, who inhabits a world at which determinism is false. This is not to say that no events are deterministically caused at Bob's world. For example, a certain radioactive particle decayed in Bob's town at noon, and that event was not deterministically caused; but the particle had been recruited as a randomizing trigger for a bomb, and once the particle decayed, the bomb was deterministically caused to explode.<sup>9</sup>

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.<sup>10</sup> The process, which is screened

---

<sup>8</sup>One of us has argued elsewhere (Mele 1995, chap. 12) that a libertarian may adopt a compatibilist conception of an action-generating process that takes an agent from a deliberative judgment about what it is best to do through a corresponding intention or decision to a corresponding action and try to locate "a theoretically useful place" (212) for indeterminism in processes leading to such a judgment.

<sup>9</sup>Even if no events are deterministically caused in the actual world, this little scenario is conceptually possible. Our concern here is with a conceptually possible agent, since *PAP* is supposed to be a conceptual truth.

<sup>10</sup>We might just as easily have imagined that  $P$  is initiated by a freak neurochemical occurrence that no agent had a role in producing. Frankfurt himself notes that there is no need for "a human manipulator," or

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, for example, he is dead by  $t_2$ ).<sup>11</sup> (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.<sup>12</sup> 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.

If this scenario is coherent, then, other things being equal (for example, Bob is sane and is not a compulsive car thief), it is plausible that Bob is morally responsible for deciding to steal the car. After all, he decided on his own to do this, with no interference

even a mechanical one; "natural forces involving no will or design at all" might do the trick (1969, n. 4).

<sup>11</sup>If Bob were to decide on his own prior to  $t_2$  to steal the car,  $P$  would cause him to decide again at  $t_2$  to do this—unless, of course, at  $t_2$ , Bob were on his own to decide again to steal the car. Notice that even in a short span of time, there are some relatively normal ways in which it might happen that a person decides twice to  $A$ . Forgetful Fred might decide at noon to call his son at 6:00, and a few minutes later, having forgotten about this, he might again decide to call his son at 6:00—or, having decided to call his daughter today to wish her a happy birthday, he might call her with birthday greetings, and then, forgetting that he called her, he might again decide to call her today to wish her a happy birthday. Cowardly Cal, having decided at noon to call for a dental appointment by the end of the business day, and having changed his mind at 12:01, might decide again at 12:02 to do this. We should mention that Ann's car is too far away for Bob to steal it by  $t_2$ . But even if it were closer and he stole it before  $t_2$ , he would decide at  $t_2$  to steal it—evidently having forgotten that he already stole it. Given the way  $P$  works, if it were to issue at  $t_2$  in Bob's deciding to steal Ann's car, it would erase any memories that are incompatible with its so issuing. (For stylistic reasons, we assume in this note that Bob is not dead or otherwise decisionally incapacitated at  $t_2$ .)

<sup>12</sup>It will become evident that a counterfactual controller's having a prior sign (Widerker) or prior knowledge (Kane) of what the agent will do is an inessential feature of Frankfurt-style cases, despite the details of Frankfurt's own case, as quoted earlier. (David Hunt independently notices this in a forthcoming article (Hunt n.d.).)

from  $P$ . And, assuming coherence, this is plausible even though, at  $t_2$ , Bob could not have done otherwise than decide to steal Ann's car. Given the details of the case, any future open to Bob after the initiation of  $P$  in which he is capable at  $t_2$  of making a decision includes his deciding at  $t_2$  to steal the car. If, in this case, Bob is morally responsible for deciding to steal Ann's car,  $PAP$  is false.

The coherence of our scenario may, however, be called into question. How, one might wonder, can it happen that Bob decides on his own at  $t_2$  to steal Ann's car, given the presence of the deterministic process we mentioned? One can understand how, *prior* to  $t_2$ , Bob might decide on his own to steal the car. (Notice that in that case, other things being equal, Bob could have decided otherwise at this earlier time, given that what he does then is causally undetermined (cf. Ginet 1996).) But how can it happen that Bob decides on his own at  $t_2$  to steal the car, and that  $P$  does not produce the decision, given what we said about  $P$ ?

Consideration of the following fanciful machine will prove useful in answering this question. The machine, designed by a specialist in machine art, produces artistic widgets of different shapes and colors. The colors of the widgets produced are determined by the color of a ball bearing (bb) that hits the machine's receptor at a relevant time. The machine,  $M$ , is surrounded by several automatic bb guns, each containing bbs of various colors. The relevant aspect of  $M$ 's mechanical design, for our purposes, is relatively simple. First, with one qualification, if a bb of color  $x$  hits  $M$ 's receptor, and  $M$  is not already in the process of making a widget,  $M$  at once starts a process designed to result in the production of an  $x$ -colored widget. Second, because two or more bbs sometimes hit the receptor simultaneously, the artist has designed his machine in such a way that whenever this happens (while  $M$  is not busy making a widget)  $M$  at once starts a process designed to result in the production of a widget the color of the right-most bb. No other striking of  $M$ 's receptor at the same time plays a role in triggering  $M$ .<sup>13</sup>

Bob is analogous to  $M$  in an important respect. He is physically and psychologically so constituted that if an unconscious deterministic process in his brain and an indeterministic decision-mak-

---

<sup>13</sup>In the interest of simplicity, we suppressed reference to another feature of the device. The rectangular receptor is too small for two bbs simultaneously to tie for the right-most position.

ing process of his were to “coincide” at the moment of decision, he would indeterministically decide on his own and the deterministic process would have no effect on his decision. This situation is an analogue of a case in which two bbs of the *same color* simultaneously hit *M*'s receptor (while *M* is not busy making a widget). Just as a blue bb's hitting the left side of the receptor at *t*, in a case in which another blue bb hits the receptor's right side at *t*, does not cause *M*'s starting a process designed to result in the production of a blue widget, *P* does not cause Bob's decision to steal Ann's car.<sup>14</sup> Instead, Bob's indeterministic deliberative process about whether to steal the car issues in his deciding to steal it, just as the striking of the right-most bb issues in *M*'s starting the widget-making process.

Cautious readers will wonder what would happen at *t2* if *P* and Bob's indeterministic deliberative process—“process *x*”—were to “diverge” at *t2*. The issue may be pictured, fancifully, as follows.<sup>15</sup> Two different “decision nodes” in Bob's brain are directly relevant. The “lighting up” of node *N1* represents his deciding to steal the car, and the “lighting up” of node *N2* represents his deciding *not* to steal the car. Under normal circumstances and in the absence of preemption, a process's “hitting” a decision node in Bob “lights up” that node. If it were to be the case both that *P* hits *N1* at *t2* and that *x* does not hit *N1* at *t2*, then *P* would light up *N1*. If both processes were to hit *N1* at *t2*, Bob's indeterministic deliberative process, *x*, would light up *N1* and *P* would not. The present question is this. What would happen if, at *t2*, *P* were to hit *N1* and *x* were to hit *N2*? That is, what would happen if the two processes were to “diverge” in this way? And *why*?

We extend Bob's story as follows. Although if both processes were to hit *N1* at *t2*, Bob's indeterministic deliberative process, *x*, would preempt *P* and light up *N1*, it is also the case that if, at *t2*, *P* were to hit *N1* and *x* were to hit *N2*, *P* would prevail. In the latter case, *P* would light up *N1* and the indeterministic process would

---

<sup>14</sup>Of course, the left-most bb's striking (idle) *M*'s receptor would have caused *M*'s starting a process for the making of a blue widget, if no bb had simultaneously struck the receptor to that bb's right, just as *P* would have caused Bob to decide at *t2* to steal the car, if Bob had not decided on his own at *t2* to steal it. (For a general defense of the possibility of direct or “occurrent” preemption, see Ehring 1997, 47–49.)

<sup>15</sup>The picture obviously is neuro-fictional, but it is useful nonetheless.

not light up *N2*. Of course, readers would like a story about why it is that although *x* would preempt *P* in the former situation, *P* would prevail over *x* in the latter. Here is one story. By *t2*, *P* has “neutralized” *N2* (but without affecting what goes on in *x*). That is why, if *x* were to hit *N2* at *t2*, *N2* would not light up.<sup>16</sup> More fully, by *t2* *P* has neutralized all of the nodes in Bob for decisions that are contrary to a decision at *t2* to steal Ann’s car (for example, a decision at *t2* not to steal anyone’s car and a decision at *t2* never to steal anything).<sup>17</sup> In convenient shorthand, by *t2* *P* has neutralized *N2* and all its “cognate decision nodes.” Bear in mind that all we need is a conceptually possible scenario, and this certainly looks like one.

The following statements (among others) are true by hypothesis in our case.

- (1) Any future open to Bob after the initiation of *P* in which he is capable at *t2* of making a decision—any future of “kind *F*”—includes his deciding at *t2* to steal the car.
- (2) In any future of kind *F*, if “an indeterministic decision-making process of Bob’s”—a process of “kind *X*”—does not issue in Bob’s deciding at *t2* to steal the car, then *P* deterministically issues in his deciding at *t2* to steal it.
- (3) In any future of kind *F*, if a process of kind *X* issues in Bob’s deciding at *t2* to steal the car, *P* does not issue in this.
- (4) A process of kind *X* does issue in Bob’s deciding at *t2* to steal the car, and *P* does not.

This collection of statements is both internally consistent and consistent with the other details of the case.

Review Bob’s indeterministic process of deliberation about whether to steal the car (process *x*). It is causally undetermined whether *x* will issue in Bob’s deciding at *t2* to steal the car, and, indeed, whether it will issue in any decision at all. Although *P* does

---

<sup>16</sup>What would happen if Bob’s indeterministic deliberative process were to hit *N2* at some time *tn* prior to *t2*? In one version of the story, *N2* would light up at *tn*—Bob would decide at *tn* not to steal the car—but then at *t2*, when *P* hits *N1*, Bob would change his mind and decide to steal it. In another version—the one we prefer, owing to its relative simplicity—*P* neutralizes *N2* as soon as Black initiates *P*.

<sup>17</sup>David Hunt independently makes a similar suggestion in a forthcoming article (Hunt n.d.).

not in any way affect what happens *in x*, *P* has neutralized *N2* and all its cognate nodes, thereby making it psychologically impossible for any indeterministic process in Bob to issue in any competing decision at all. It is possible that Bob's indeterministic process, *x*, will indeterministically issue in his deciding at *t2* to steal the car, without *P* in any way contributing to this, and it is possible that *P* will issue in his deciding at *t2* to steal the car (which will happen if Bob does not indeterministically decide on his own at *t2* to steal it). But, given *P*'s neutralizing effects, it is not possible that *x* or any other indeterministic process in Bob will issue in his making a decision at *t2* that is incompatible with his deciding at *t2* to steal the car. This is so even though *P* does not influence what happens in *x* itself. Furthermore, Bob has all along been psychologically so constituted that *x* would preempt *P* were *x* to hit *N1* (the node for deciding to steal the car) at *t2*. Bob's psychological constitution is an analogue of machine *M* in that connection.

One might object that given a nomic subsumption model of causality (*NSM*), if Bob's deciding at *t2* to steal Ann's car is entailed by a correct description of some state of the world prior to *t2* and the laws of nature, that suffices for Bob's decision's being deterministically caused, appearances to the contrary notwithstanding.<sup>18</sup> If Bob's decision to steal the car is deterministically caused, we face a criticism that Kane and Widerker leveled against Frankfurt's own scenarios. Now, our imagined opponent may have either of two views of *NSM* in mind. We consider them in turn. (1) If *NSM* asserts that whenever *a* causes *b*, there is a law subsuming them, then the subsumption model does not entail that *P* (the deterministic process initiated by Black) or anything else caused—deterministically or otherwise—Bob's decision. Moreover, it is compatible with this version of *NSM* that *x*, Bob's indeterministic decision-making process, indeterministically caused the decision, since *x* and Bob's decision may be subsumed by a probabilistic law. (2) If *NSM* asserts that whenever there is a law subsuming *a* and *b*, *a* causes *b*, then this might seem worrisome, especially if, according to the law, *a* makes *b* inevitable. Notice that this version of *NSM* is plausible only if 'law' here means 'causal law', and a causal law that is to be prob-

---

<sup>18</sup>Fischer considers an objection of this kind to standard Frankfurt-style cases (1982, 35–37). Several readers of an earlier draft urged us to address it.

lematic for us cannot be of the form “Events of *a*’s type cause events of *b*’s type unless the former are preempted.” However, the law subsuming *P* and Bob’s decision is in fact a law to the effect that processes of *P*’s type will deterministically cause decisions of the pertinent type unless the former are preempted, and this law does not entail that *P* or anything else deterministically caused the decision.<sup>19</sup>

Perhaps the intended objection is not that *P* deterministically caused the decision, but just that the decision was *determined*. If ‘determined’ means ‘inevitable’, that is fine with us. After all, the primary thrust of Frankfurt’s own case is that the agent is morally responsible for his *A*-ing even though his *A*-ing is inevitable, and the point of our example is that Bob is morally responsible for deciding (at *t*<sub>2</sub>) to steal the car even though he could not have done otherwise (then) than decide to steal it.<sup>20</sup> However, if ‘determined’ means ‘deterministically caused’, we are entitled to ask *what* deterministically caused the decision. It was not *P*. So was it the conjunction (the mereological sum) of *P* and *x*, or of *P*, *x*, and various other things? Even supposing that a gerrymandered entity like this can be an appropriate candidate for a cause, it is difficult to see how it can have deterministically caused Bob’s decision, given that *P* did *not* cause the decision and that *x* caused it only indeterministically.

Consider matters from an intuitively appealing perspective (cf. Fischer 1982, 33).<sup>21</sup> Subtract Black and *P* from our scenario and imagine that what happens at Bob’s indeterministic world is that *x*, Bob’s indeterministic decision-making process, indeterministically issues at *t*<sub>2</sub>—in some way favored by libertarians—in his deciding to steal the car. Plainly, there is no deterministic cause of

---

<sup>19</sup>It may be claimed that there are no “unless preempted” laws of this kind. But this requires an argument that goes well beyond a mere appeal to *NSM*.

<sup>20</sup>Strictly speaking, there is no assurance that Bob’s deciding at *t*<sub>2</sub> to steal the car is *inevitable* in our scenario. It is consistent with the background details of the scenario, for example, that Bob dies just as he is about to decide to steal the car, owing to a causally undetermined lightning strike. As we put it earlier, “Given the details of the case, any future open to Bob . . . *in which he is capable at t<sub>2</sub> of making a decision* includes his deciding at *t*<sub>2</sub> to steal the car” (emphasis added). However, we deemed it best to meet head on the objection introduced in the preceding paragraph.

<sup>21</sup>John Fischer commended this perspective to us in correspondence.

Bob's decision in this case. Now add Black and *P* to the scenario in just the way we have done. At *t*<sub>2</sub>, process *x* issues in the same indeterministic way in Bob's decision: by hypothesis, Black and *P* do not influence *x*. Although at *t*<sub>2</sub> Bob cannot do otherwise than decide then to steal the car, nothing warrants the claim that his decision is deterministically caused.

Bob's case is coherent, and it apparently falsifies *PAP*. Again, other things being equal, Bob certainly seems to be morally responsible for deciding to steal Ann's car, for he decided on his own to steal it. Even so, he could not have done otherwise than decide to steal it, and he could not have done otherwise at *t*<sub>2</sub> than decide to steal it.<sup>22</sup> And notice that the case is what Widerker calls an "*IRR*-situation": there are "circumstances in which" Bob decides to steal Ann's car that "make it impossible for him to avoid" deciding to do this but "in no way bring it about that" he decides to do this (1995a, 248). Although process *P* makes it impossible for Bob to avoid deciding to steal Ann's car, *P* in no way brings it about that he decides to do this. Instead, Bob indeterministically decides on his own at *t*<sub>2</sub> to steal the car.

There are ways to resist our conclusion. For example, one can argue that although Bob is morally responsible for deciding *on his own* to steal the car, he is not morally responsible for deciding to steal it (Naylor 1984). (Even though Bob could not have done otherwise than decide to steal the car, he could have done otherwise than decide on his own to steal it: he could have decided to steal it because *P* made him decide this.) Alternatively, one can contend that for the proponent of *PAP*, 'he could have done otherwise' in that principle simply means 'he was not causally determined to do what he did', in which case the proponent can consistently grant that Bob is morally responsible for deciding to steal the car (Heinaman 1986, 275–76). These lines of response to Frankfurt-style scenarios have been in circulation for some time; an adequate rebuttal of them would require a paper of its own and would duplicate the efforts of others.<sup>23</sup> Our concern is with the more recent and potentially more damaging criticisms of Frank-

---

<sup>22</sup>Again, this is consistent with its being the case that something else could have *happened* than that Bob decided to steal the car. For example, Bob might have died before *t*<sub>2</sub>.

<sup>23</sup>For a careful critical discussion of various incompatibilist responses to Frankfurt-style cases, see Fischer 1994, chap. 7.

furt-style scenarios by Widerker and Kane.<sup>24</sup> Our little story about Bob undermines those criticisms.<sup>25</sup>

### 3. A Historical Version of *PAP*

Some incompatibilists favor a historical version of *PAP* along the following lines:

*PAPh*: *S* is morally responsible for what he did at *t* only if (1) he could have done otherwise at *t* or (2) even though he could not have done otherwise at *t*, the psychological character on the basis of which he acted at *t* is itself partially a product of an earlier action (or actions) of his performed at a time when he could have done otherwise.<sup>26</sup>

If our case falsifies *PAP*, an expanded version of the case falsifies *PAPh*. What we have in mind now is what one of us elsewhere

<sup>24</sup>Naylor's and Heinaman's responses each make a significant concession to the Frankfurt-style argument. Naylor concedes that the agent could not have done otherwise than *A*. To be sure, she contends that the agent is not morally responsible for *A*-ing, but rather for *A*-ing on his own (cf. van Inwagen 1983, 181). However, that contention has been deemed an instance of biting an intolerably hard bullet even by a libertarian (Kane 1996, 41–42; see Ginet 1996, 407 for another criticism). Heinaman concedes that the agent could not have done otherwise in the sense that he did not have open to him at the time the possibility of acting otherwise. (We have a bit more to say about his response later.) The arguments by Widerker and Kane that we have sketched make no concessions to the Frankfurt-style argument.

<sup>25</sup>It also avoids Wyma's and Ginet's criticisms. Wyma's complaint about Frankfurt-style cases (1997, 63–67) is essentially Widerker's. Ginet contends that if an agent *S* cannot do otherwise than *A* by a time *t*\* in a Frankfurt-style case, then *S* is not morally responsible for his *A*-ing by *t*\*, and he argues that, even so, *S* may be morally responsible for *A*-ing at the earlier time, *t*, at which he in fact *A*-ed—a time at which *S* was able to do otherwise than *A* then (1996, 406–7). If Ginet is right, *PAP* survives the case he imagines. On his view, what *S* is morally responsible for, if anything, is a state of affairs that he could have avoided: namely, its being the case that he *A*-s at *t*. However, in our scenario, at *t*2, the very time at which Bob decides to steal Ann's car, he cannot do otherwise than decide to steal it, and, even so, it is plausible, for reasons that we have set out, that he is morally responsible for the decision he made at *t*2 and for his so deciding at that time.

<sup>26</sup>Kane advances a view of this kind (1996, 39–43). Mele, who is not a libertarian, has suggested that libertarians should prefer a historical principle of this kind to *PAP* (1995, 208–9).

termed a *global* Frankfurt-style case—a case in which at any relevant earlier time the agent acts “on his own” but could not have done otherwise, owing to a circumstance of the sort mentioned in *IRR*.<sup>27</sup>

Imagine, for example, that in *any* case in which Bob makes a decision—a decision to *A*, say—a deterministic process like *P* was under way that would have resulted in his deciding to *A* if he had not indeterministically decided on his own to *A*, and imagine that the same thing is true also of all of Bob’s actions that are not decisions. This global fact about Bob is quite remarkable, but it is a coincidence nonetheless. Given this fact, Bob could never have done otherwise than he did. But since he did everything on his own, the deterministic processes always having been preempted, we see no good reason to hold that the presence of those deterministic processes deprives him of moral responsibility.

It should be emphasized that we ourselves are not inclined to use our new Frankfurt-style case in either its local or its global incarnation as part of an argument for compatibilism.<sup>28</sup> After all, there is a significant difference between our scenarios and scenarios set in deterministic worlds: in ours, as is supposed to be the case in standard Frankfurt-style scenarios in general, the agent’s decisions lack deterministic causes. In principle, a libertarian’s incompatibilism might be motivated, not by the thought that determinism precludes our ever having been able to do otherwise than we did, but instead by the thought that in a deterministic world our actions (including our decisions) are ultimately causally ensured *consequences* of the laws of nature and states that obtained long before we were born.<sup>29</sup> Indeed, one of the responses to Frankfurt-style scenarios that we mentioned in passing—the response that in *PAP*, ‘he could have done otherwise’ simply means ‘he was not causally determined to do what he did’—suggests that the latter thought might be the real worry for some incompatibilists.

Those who took arguments of the sort offered by Widerker and

---

<sup>27</sup>See Mele 1996, 129. Cf. Fischer 1994, 214; Mele 1995, 141; Haji 1996, 707; Kane 1996, 42–43, 143; Wyma 1997, 67.

<sup>28</sup>Mele is officially agnostic about the main metaphysical issue that separates compatibilists from incompatibilists (1995).

<sup>29</sup>The latter thought is featured in the incompatibilist “consequence argument.” For a short version of the argument, see van Inwagen 1983, 16. For detailed versions, see, for example, Ginet 1990, chap. 5 and van Inwagen, chap. 3.

Kane to show that incompatibilists can safely dismiss all Frankfurt-style cases should start worrying again. In rescuing the Frankfurt-style scenario, we are not directly attacking incompatibilism. However, we note that it is incumbent upon libertarians and other incompatibilists to explain what about determinism is incompatible with moral responsibility (and the freedom required for such responsibility), if not that determinism is inconsistent with our having been able to do otherwise than we did.<sup>30</sup>

*Davidson College (Mele)*

*Brooklyn College, CUNY (Robb)*

#### References

- Ehring, Douglas. 1997. *Causation and Persistence*. Oxford: Oxford University Press.
- Fischer, John. 1982. "Responsibility and Control." *Journal of Philosophy* 79: 24–40.
- . 1994. *The Metaphysics of Free Will*. Oxford: Blackwell.
- . 1995. "Libertarianism and Avoidability: A Reply to Widerker." *Faith and Philosophy* 12:119–25.
- Fischer, John, and Paul Hoffman. 1994. "Alternative Possibilities: A Reply to Lamb." *Journal of Philosophy* 91:321–26.
- Frankfurt, Harry. 1969. "Alternate Possibilities and Moral Responsibility." *Journal of Philosophy* 66:829–39.
- Ginet, Carl. 1990. *On Action*. Cambridge: Cambridge University Press.
- . 1996. "In Defense of the Principle of Alternative Possibilities: Why I Don't Find Frankfurt's Argument Convincing." *Philosophical Perspectives* 10:403–17.
- Haji, Ishtiyaque. 1996. "Moral Responsibility and the Problem of Induced Pro-Attitudes." *Dialogue* 35:703–20.
- Heinaman, Robert. 1986. "Incompatibilism Without the Principle of Alternative Possibilities." *Australasian Journal of Philosophy* 64:266–76.
- Hunt, David. n.d. "Moral Responsibility and Unavoidable Action." *Philosophical Studies*, forthcoming.
- Kane, Robert. 1985. *Free Will and Values*. Albany: State University of New York Press.
- . 1996. *The Significance of Free Will*. New York: Oxford University Press.

---

<sup>30</sup>For a relevant suggestion, see Mele 1996. For discussion or comments, we are grateful to Laura Ekstrom, John Fischer, Carl Ginet, Bob Kane, Hugh McCann, and Tim O'Connor.

- Lamb, James. 1993. "Evaluative Compatibilism and the Principle of Alternate Possibilities." *Journal of Philosophy* 90:517–27.
- Mele, Alfred. 1995. *Autonomous Agents*. New York: Oxford University Press.
- . 1996. "Soft Libertarianism and Frankfurt-Style Scenarios." *Philosophical Topics* 24:123–41.
- Naylor, Margery. 1984. "Frankfurt on the Principle of Alternate Possibilities." *Philosophical Studies* 46:249–58.
- Stump, Eleonore. 1990. "Intellect, Will and the Principle of Alternate Possibilities." In *Christian Theism and the Problems of Philosophy*, ed. Michael Bealy, 254–85. Notre Dame: University of Notre Dame Press.
- Stump, Eleonore, and Norman Kretzmann. 1991. "Prophecy, Past Truth, and Eternity." *Philosophical Perspectives* 5:395–424.
- Van Inwagen, Peter. 1983. *An Essay on Free Will*. Oxford: Oxford University Press, Clarendon Press.
- Widerker, David. 1995a. "Libertarianism and Frankfurt's Attack on the Principle of Alternative Possibilities." *Philosophical Review* 104:247–61.
- . 1995b. "Libertarian Freedom and the Avoidability of Decisions." *Faith and Philosophy* 12:113–18.
- Wyma, Keith. 1997. "Moral Responsibility and Leeway for Action." *American Philosophical Quarterly* 34:57–70.
- Zagzebski, Linda. 1991. *The Dilemma of Freedom and Foreknowledge*. Oxford: Oxford University Press.