
THE JOURNAL OF PHILOSOPHY

VOLUME CXIII, NO. 11, NOVEMBER 2016

PAP-STYLE CASES*

Over the years, two models of freedom have emerged as competitors. (My focus is metaphysical freedom, understood as the metaphysical condition of moral responsibility, or the kind of control that is required to be morally responsible. In this paper, I examine certain claims about responsibility, but mainly insofar as they concern the metaphysical condition.) On the one hand, there is the *alternative-possibilities* model, which states that acting freely consists, at least partly, in being able to do otherwise. On the other hand, there is the *actual-sequence* model, which states that acting freely is exclusively a function of the actual sequence of events issuing in our behavior; on this view, alternative possibilities are irrelevant to freedom.¹

The alternative-possibilities model, once the dominant view, was famously challenged by the introduction of *Frankfurt-style cases*.² In a Frankfurt-style case, an agent appears to act freely and be morally

* Thanks to Juan Comesaña, Michael McKenna, and two anonymous reviewers. Thanks also to audiences at the Center for the Philosophy of Freedom at the University of Arizona and at a workshop on agency, responsibility, and character at the Universidad del Rosario and the Universidad de Los Andes in Bogotá.

¹ For a classic development of the alternative-possibilities model, see Peter van Inwagen, *An Essay on Free Will* (New York: Oxford University Press, 1983). For a classic development of the actual-sequence model, see John Martin Fischer and Mark Ravizza, *Responsibility and Control: A Theory of Moral Responsibility* (Cambridge, UK: Cambridge University Press, 1998). Although Fischer and Ravizza sometimes advertise their theory as a theory of moral responsibility, it is in fact a theory of the freedom or control condition only, as they note in the introductory chapter (see pp. 13–14). The claim that alternative possibilities are irrelevant to freedom follows from the claim that freedom is exclusively a function of actual sequences, given certain plausible assumptions about the constitution of actual sequences. I discuss this in *Causation and Free Will* (Oxford: Oxford University Press, 2016), chapter 1.

² These cases were originally developed by Harry Frankfurt in his “Alternate Possibilities and Moral Responsibility,” this JOURNAL, LXVI, 23 (December 1969): 829–39.

responsible for an action that he or she could not have avoided. The agent could not have avoided the action because, as Frankfurt imagined it, a neuroscientist who had been secretly monitoring the agent's brain activity would not have let him make a different choice (for example, if he had shown any signs that he was about to make a different choice, the neuroscientist would have forced him to make the same choice via direct manipulation of his brain). Still, the agent seems responsible for what he does and seems to act freely, for he chooses to act completely on his own, on the basis of his own reasons. Hence, Frankfurt-style cases seem to show that the kind of freedom relevant to responsibility does not require the ability to do otherwise but is instead just a function of the factors that actually account for the agent's behavior. Thus, they seem to challenge the alternative-possibilities model and support the actual-sequence model. Whether they are actually successful in doing this is still a matter of debate; in fact, a large part of the recent literature on free will has focused on this question.³

Here I will examine a related but different question. To motivate it, let me start by noting, again, that even if there was a time when the alternative-possibilities view was the dominant view of freedom, matters are different now: these days, the actual-sequence model is regarded as sufficiently plausible on its own to stand as a serious competitor. Arguably, this is the case independently of what one thinks about the strategy of appealing to Frankfurt-style cases to refute the alternative-possibilities model. For, even assuming that this strategy could never be made to work, the central idea behind the actual-sequence view seems independently plausible. As Frankfurt noted when he introduced the Frankfurt-style cases, it is in fact quite intuitive to think that the kind of freedom relevant to responsibility is only a function of the factors that actually explain our behavior. In particular, it would seem inappropriate for agents to try to absolve themselves of responsibility for actions they performed by appealing to factors that did not at all account for what they did and that played no role whatsoever in determining the actual sequence of events issuing in their behavior (for example, it would seem inappropriate for an agent in a Frankfurt-style case who then becomes aware of the

³For a collection of works on this topic, see David Widerker and Michael McKenna, *Moral Responsibility and Alternative Possibilities: Essays on the Importance of Alternative Possibilities* (Aldershot: Ashgate, 2003). Although Frankfurt's main concern was moral responsibility, his argument can easily be reformulated in terms of freedom (the metaphysical condition of responsibility), so that is how I am presenting it here.

neuroscientist's existence to try to excuse his behavior by drawing attention to the presence of the neuroscientist).⁴

So there are two competing models of freedom, and they each have their own share of intuitive appeal. How are we to find out which one is true? Identifying what freedom is or what it requires is an important project in its own right, but it also has significant implications for some central free will debates. In particular, if one believes that determinism is incompatible with the existence of alternative possibilities (if determinism is true, we could never have done otherwise), then the question of whether freedom requires alternative possibilities or is instead just a function of the actual sequence can be key to establishing whether *compatibilism* is true—that is, whether we can be free in a deterministic world.

In general, a natural strategy when trying to decide between two models of a certain concept is to look for examples that appear to refute one model and support the other. Again, Frankfurt-style cases have been used for this purpose, to challenge the alternative-possibilities view and support the actual-sequence view; some have argued that they succeed in doing this, others that they fail. But how about examples that seem to show the *opposite*—that is, that the actual-sequence view is false and the alternative-possibilities view is true? Are there any such examples? What do they look like, and are they successful?

That is the topic of this paper. It is an issue that has received very little attention in the literature. However, it seems to be extremely important if one is to get a better perspective on the whole debate. For then, arguably, if both views are independently plausible, assessing the prospects of the counterparts of Frankfurt-style cases is as important as assessing the prospects of Frankfurt-style cases themselves. Here I will not take a stand on the issue of whether Frankfurt-style cases are successful, but I will argue that their counterparts fail, and that they fail for a principled reason. Thus, if one is interested in defending the actual-sequence view, this is a significant step in that direction.

⁴ See Frankfurt, "Alternate Possibilities and Moral Responsibility," *op. cit.*, section v. See also Harry Frankfurt, "Some Thoughts Concerning PAP," in Widerker and McKenna, eds., *Moral Responsibility and Alternative Possibilities*, *op. cit.*, pp. 339–45; Michael McKenna, "Frankfurt's Argument against Alternative Possibilities: Looking Beyond the Examples," *Noûs*, XLII, 4 (December 2008): 770–93; and Linda Zagzebski, "Does Libertarian Freedom Require Alternate Possibilities?," *Philosophical Perspectives*, xiv (2000): 231–48. I discuss the motivation for the actual-sequence view and, in particular, how Frankfurt-style cases can be put to use in articulating some of that motivation, in "Frankfurt-Style Examples," in M. Griffith, N. Levy, and K. Timpe, eds., *Routledge Companion to Free Will* (London: Routledge, 2017), pp. 179–90.

I will call the counterparts of Frankfurt-style cases *PAP-style cases*. (“PAP” is the standard acronym for the “principle of alternative possibilities,” the claim that is at the heart of the alternative-possibilities model and that states that freedom requires access to alternative possibilities.⁵) A PAP-style case, then, is an example that seems to show that freedom is not exclusively a function of actual sequences but rather requires access to alternative possibilities.

Arguably, there is an important structural difference between Frankfurt-style cases and PAP-style cases. This is due to a structural difference between the alternative-possibilities model and the actual-sequence model. As explained above, the alternative-possibilities model aims to identify a *necessary* condition for freedom (one that is, of course, substantial or robust enough; perhaps one that can help ground freedom). In contrast, the actual-sequence model does not mainly aim to identify a necessary condition for freedom; if anything, it aims to *rule out* a certain condition as a necessary condition by claiming that other types of factors different from alternative possibilities are *sufficient* grounds for freedom. So whereas Frankfurt-style cases aim to show that having alternative possibilities is not a necessary condition for freedom, PAP-style cases aim to show that it is.

Now consider: What would it take to show, by appeal to examples, that something is *not* a necessary condition for freedom? Clearly, it would take a single example in which an agent acts freely in the absence of that condition (this is the structure of a Frankfurt-style case). In contrast, consider: What would it take to show, by appeal to examples, that something *is* a necessary condition for freedom? Here a single example is not enough. (What would that scenario be? It cannot be one in which an agent acts freely and has alternative possibilities, and it also cannot be one in which an agent does not act freely and lacks alternative possibilities; these examples cannot by themselves establish that freedom requires alternative possibilities.) Presumably, one should instead look for a *pair* of closely related contrast scenarios that suggest that having those alternatives can *make the difference* between acting freely and not acting freely, and thus between being responsible and not being responsible. Arguably, these would have to be, first, a scenario where an agent is intuitively responsible and has alternative possibilities, and, second, a scenario where he or she is intuitively not responsible and lacks alternative possibilities *but is otherwise just like the first scenario* (in all potentially relevant respects). If

⁵ See, notably, van Inwagen, *An Essay on Free Will*, *op. cit.*; and Carl Ginet, *On Action* (Cambridge, UK: Cambridge University Press, 1990).

there are such scenarios, they can be used to argue for the alternative-possibilities view and against the actual-sequence view.⁶

Are there such scenarios? I turn to that question in the next section.

I

Before discussing what I take to be the best candidates for PAP-style cases, let me briefly discuss a class of poorer candidates and explain why they fail. Doing so will help identify some important features of the better candidates.

When Frankfurt presented his challenge to the alternative-possibilities model, he also attempted to explain why that model had seemed so appealing to many people, even though it is in fact false. He suggested that its appeal was the result of a natural conflation of two kinds of factors that can be operative in certain cases: the factors that make an action unavoidable, and the factors that actually account for an action.⁷ I will use the labels ‘inevitability factors’ and ‘explanatory factors’, respectively.

As Frankfurt noted, inevitability factors and explanatory factors tend to go hand in hand: when an agent’s action is unavoidable, and thus there are inevitability factors that make it unavoidable, they *also* tend to be the explanatory factors. For example, in a scenario involving coercion, the coercive threat both makes the act unavoidable and accounts for the agent’s behavior: the threat is what causes the agent’s behavior. In a scenario of that kind, the agent does not act freely. However, Frankfurt argued that the inevitability factors and explanatory factors can come apart and that, when they do, we see how agents can act freely in the absence of alternative possibilities. For example, in a Frankfurt-style case, the inevitability factors are not the same as the explanatory factors (the inevitability factors concern the neuroscientist, and the explanatory factors concern the agent’s own deliberation on the basis of his own reasons), so the agent can act freely even when he is unable to do otherwise. According to Frankfurt, what determines whether agents have the relevant kind of freedom are just the explanatory factors. Still, the fact that the inevitability factors tend to also be the explanatory factors when agents are unable to do otherwise accounts for why it seemed so natural to think that those agents cannot act freely.

Regardless of whether Frankfurt is right about this, the distinction between the inevitability factors and the explanatory factors is

⁶Here I will set aside potential problems with the general methodology of using contrast scenarios identified by Shelly Kagan in “The Additive Fallacy,” *Ethics*, xcix, 1 (October 1988): 5–31. I am happy to grant that those problems do not arise in this case.

⁷Frankfurt, “Alternate Possibilities and Moral Responsibility,” *op. cit.*, sections I and II.

important for our purposes in that it can help us see more clearly the kind of structure that a PAP-style case would have to have. Recall that a PAP-style case is actually a pair of scenarios: one scenario where an agent is intuitively responsible and has alternative possibilities, and another scenario where the agent is intuitively not responsible and lacks alternative possibilities but is otherwise the same (in all potentially relevant respects) as the first one. So now imagine that we wanted to use a coercion scenario and a scenario not involving coercion as the two cases in the pair. This would not work. Given that in a coercion scenario the inevitability factors are *also* the explanatory factors, they enter the actual sequence of events issuing in the agent's behavior. Thus, any scenario not involving coercion, regardless of how similar it might be in other respects, would differ from it with respect to the actual sequence of events and not just with respect to the agent's ability to do otherwise. Thus, the two scenarios will fail to be alike in all potentially relevant respects. In particular, the actual-sequence model would have an easy time explaining the difference in responsibility between the two cases without appealing to the difference in alternative possibilities simply by drawing attention to the difference in the actual sequence.

The lesson to be learned from this is that a PAP-style case cannot include scenarios in which the inevitability factors are also the explanatory factors. In other words, the inevitability factors must not themselves be part of the actual sequence of events. This reduces considerably the class of potential candidates.

Still, there are several examples in the literature that might be thought to fill the bill. Some of them were offered by advocates of the alternative-possibilities model as part of a response to Frankfurt's challenge.⁸ As we will see, they all have the same basic structure.

Let us start with an example by Carl Ginet. Ginet once motivated the alternative-possibilities model by appealing to the following elegant scenario.⁹ Imagine that I think that I could turn on the light by flipping a switch, but I decide against it. Imagine that I was wrong and I could not have turned the light on, for the bulb was burned out. So, in fact, I had no real alternatives (with respect to the light being off); I only thought that I did, but that belief was an illusion. In that case, it seems plausible to think that I am not responsible for the light staying off because I lack the relevant kind of freedom or control with respect to that state of affairs due to my lack of alternatives.

⁸ van Inwagen, in particular, is explicitly responding to Frankfurt when he presents the examples discussed below (see van Inwagen, *An Essay on Free Will*, *op. cit.*, chapter V).

⁹ Ginet, *On Action*, *op. cit.*, p. 91.

As explained above, it will help to think about this case as a contrast between two scenarios:

Light Bulb 1: The bulb was not burned out. I am responsible for the light staying off.

Light Bulb 2: The bulb was burned out. I am not responsible for the light staying off.

Whereas in Light Bulb 1 I have alternative possibilities (with respect to the relevant state of affairs), in Light Bulb 2, I do not. This is a difference between the two cases. If it were the *only* difference, or the only potentially relevant one, then this would suggest that the difference in responsibility is due to the difference in alternative possibilities. And in this case (unlike in the previous cases involving coercion and the absence of coercion) it does seem quite plausible to argue that it is the only potentially relevant difference. In particular, given that I never even try to turn off the light, the light bulb being burned out (or the filament being broken) in Light Bulb 2 is not part of the actual sequence of events. In this respect, it is unlike a scenario of coercion, in which the coercive threat is part of the actual sequence: here the inevitability factors are *not* the same as the explanatory factors, the actual sequence seems to be the same in both cases, and the only potentially relevant difference between them seems to be the existence of alternative possibilities. For this reason, the pair of light bulb scenarios appears to be a good candidate for a PAP-style case.

Next, let us consider a couple of examples by Peter van Inwagen.¹⁰ I will cut to the chase and present them together, each of them as a pair of scenarios:

Phones: I witness a man being robbed and beaten. I consider calling the police. I could easily pick up the phone and call them, but I decide against it. I am responsible for not calling the police.

No Phones: Unbeknownst to me, the phone lines were down at the time, so I could not have called the police. I am not responsible for not calling the police.

Not All Roads Lead to Rome: A man, Ryder, is riding a runaway horse, Dobbin. Ryder cannot stop Dobbin, but he can steer him in different directions by using his bridle. Ryder hates Romans and wants them to get hurt. So, when they approach a certain crossroad, he steers Dobbin in the direction that he knows is the only path leading to Rome. As predicted, some Romans get hurt by the horse. Ryder is responsible for the fact that some Romans are hurt.

¹⁰van Inwagen, *An Essay on Free Will*, *op. cit.*, sections 5.4 and 5.6.

All Roads Lead to Rome: Ryder was wrong: in fact, all roads led to Rome. He is not responsible for the fact that some Romans are hurt.

The same is true of these cases. There is a difference in the agent's responsibility between the scenarios in each pair, as well as a difference in the existence of alternative possibilities. Moreover, the difference in alternative possibilities appears to be the only potentially relevant difference in each case. In particular, when the agent could not have done otherwise, the factors that would have prevented him from doing otherwise are not part of the actual sequence, for the agent never tries to do otherwise (that is, the broken phone lines never stop me from calling the police, for I never try to call them, and Dobbin never gets to run on the other paths that lead to Rome, for Ryder does not steer him in those other directions). So there seems to be no difference in the actual sequence.

Finally, let us consider a scenario discussed by John Martin Fischer and Mark Ravizza.¹¹ Again, it helps to think about it as a pair of related scenarios:

No Sharks: An agent is walking along the seashore when he sees a child drowning and crying for help. He could easily rescue the child. Unmoved by the situation, he decides not to intervene, and the child drowns. The agent is responsible for failing to save the child, and for the child's death.

Sharks: Unbeknownst to the agent, he could not have rescued the child. If he had tried to rescue him, some hungry sharks would have attacked him and prevented him from saving the child. The agent is not responsible for failing to save the child, or for the child's death.

Once again, there is a difference in responsibility and a difference in alternative possibilities, and this seems to be the only potentially relevant difference between the scenarios. In particular, given that the agent never tried to save the child, the presence of the sharks in *Sharks* makes no difference to the actual sequence, which seems to be the same as in *No Sharks*.

We have now identified the structure of a PAP-style case and several possible examples with that structure. As we have seen, a PAP-style case is a pair of closely related scenarios, one scenario in which the agent is intuitively responsible and has alternative possibilities and another scenario in which the agent is intuitively not responsible and lacks alternative possibilities; in addition, the only potentially relevant difference between the two scenarios is the existence of alternative possibilities. In particular, in the scenario where the agent lacks

¹¹ Fischer and Ravizza, *Responsibility and Control*, *op. cit.*, p. 125.

alternative possibilities, the inevitability factors (the factors by virtue of which the agent lacks alternative possibilities) do not account for the agent's behavior, so the actual sequence appears to be the same in both scenarios.

A striking feature of these scenarios is that most of them involve omissions by agents (van Inwagen's Ryder and Dobbin cases are the only ones that do not). It is natural to wonder why. I believe that this feature is telling, and I will return to it momentarily. For now, I will focus on the question of whether the scenarios are successful as PAP-style cases.

II

I will argue that the examples fail, and that they fail for a principled reason that is the same reason in all cases. I will illustrate this with the No Sharks/Sharks pair. (As we have seen, the examples are structurally similar. I will briefly explain how to extend the argument to the other cases at the end of this section.) I will argue that the examples fail because the difference in alternative possibilities is *not* the only potentially relevant difference between the members of each pair; in fact, there is always a difference in the *actual sequence* of events—the same kind of difference in all cases, and one that is intuitively relevant.

Let me start by drawing attention to an important feature of the No Sharks/Sharks scenarios. In both scenarios, the agent is intuitively blameworthy for some things, even in the case where the sharks are present: he is blameworthy, in particular, for deciding not to attempt a rescue. In both cases, the agent believes that he could easily save the child but remains unmoved. Also, his decision not to attempt a rescue is not coerced in any way but is made freely. So what leads him to decide not to attempt a rescue is exactly the same in both cases, and it was in fact open for him to make a different decision. Thus, the agent in Sharks is responsible for his decision; what he is *not* responsible for is the child's death, or his failure to save the child.

This is important because, again, for a pair of scenarios to be a good candidate for a PAP-style case, the *only* potentially relevant difference between them must be the existence of alternative possibilities. If the agent had made the decision for different reasons in each case, then the difference in alternative possibilities would *not* be the only potentially relevant difference between the cases; there would also be a difference in the actual sequence issuing in the decision, and then that difference could account for the difference in responsibility. So the difference in responsibility between No Sharks and Sharks concerns not the agent's responsibility for certain decisions that he made but rather his responsibility for certain *results*. In Sharks, the agent is not responsible for

the child's death, or for his failure to save the child; in No Sharks, he is responsible for both of these things. The child's death is a consequence in the world, and the agent's failure to save the child is an omission that involves a consequence in the world (the child's death, or the agent's not saving the child). So both involve results, as opposed to mere decisions.¹²

Now, consider how agents can become responsible for results. When agents are responsible for results, they are not responsible for them directly but rather become responsible for them by being responsible for other things that lead to those results. To take a simple example, an assassin who shoots his victim can be responsible for his victim's death by being responsible for his decision to shoot: his responsibility for the decision carries over to the outcome in the world. But, why does responsibility carry over in this way, when it does? Presumably, this reasoning is captured by a principle of this kind:

Principle of Transmission of Responsibility (PTR): If an agent is responsible for X , X causes Y , and the agent satisfies the relevant epistemic conditions for responsibility, then the agent is also responsible for Y .

I will work with a specific formulation of this principle. First, I will restrict responsibility to blameworthiness, as this is what is at stake in the examples reviewed. Second, I will assume that the epistemic conditions are satisfied in circumstances where part of the reason that the agent is blameworthy for X is precisely that he could foresee that X was likely to cause Y , in roughly the way that it did. Whereas it is hard to say generally what the epistemic conditions for responsibility are, it is very reasonable to think that they are met in circumstances of this kind.

Note how plausible this principle seems, especially under that specific interpretation. In particular, note that the principle offers a very natural account of the way in which the assassin becomes blameworthy for his victim's death: he is blameworthy for his victim's death because he is blameworthy for his decision to shoot, the decision to shoot causally resulted in his victim's death, and he is blameworthy for his decision to shoot partly because he could foresee that it was likely to result in his victim's death, in roughly the way that it did.

I will argue that PTR can be used to show that the No Sharks/Sharks pair fails as a PAP-style case. In particular, it can be used to argue that the actual sequence is not the same in Sharks and No Sharks;

¹² But couldn't there be pairs of scenarios where the actual sequence issuing in the decision is the same, even though in one case the agent has alternative possibilities with respect to the decision and in the other case he does not? Couldn't *they* work as PAP-style cases? I discuss this question in section III.

therefore, the existence of alternative possibilities is not the only potentially relevant difference between the two scenarios.¹³

First, a note of clarification is in order. As I mentioned above, most of the examples discussed (including the one I chose to focus on, the No Sharks/Sharks pair) involve omissions. For simplicity, I will work under the assumption that omissions can be causes. This assumption is needed in order to be able to apply PTR to cases involving omissions, for as stated, the principle includes a causal condition. But it is important to realize that the assumption is not, strictly speaking, necessary and that the heart of the argument is consistent with different views on omissions and their causal powers. For even those who believe that omissions cannot be causes tend to believe that omissions can make us morally responsible for things. Consider, for example, Dowe's view: Dowe argues that omissions cannot be causes but only "quasi-causes," where quasi-causation is, essentially, merely possible causation (for example, the non-watering of the plant quasi-causes its death because watering it *would* have caused it to continue on living).¹⁴ On this view, my failing to water the plant can make me responsible for its death in light of the existence of this quasi-causal relation. This means that a view of this kind would have to understand PTR in a different way (when applied to omissions), as making reference to quasi-causation instead of causation.

For similar reasons, my argument does not hinge on any specific ontological view of omissions; in particular, it does not hinge on a realist view according to which absences are genuine entities. My failing to water the plant could make me responsible for its death, even if it is not a real entity and thus not one of its causes, because, for example, making reference to the fact that I failed to water the plant could convey information that is explanatorily relevant to the plant's death, by telling us something about its causal history (that is, it does not contain a watering episode of some kind) and about the causal structure of nearby possible worlds (that is, in worlds where I water the plant, the plant lives as a result).¹⁵ Here, too, we have responsibility-grounding

¹³In *Causation and Free Will*, *op. cit.*, chapter 2, I give a similar argument for the claim that cases of this kind fail to undermine a certain *supervenience* principle that (I argue) is central to actual-sequence views of freedom—roughly, the claim that freedom supervenes on actual sequences.

¹⁴See Phil Dowe, "A Counterfactual Theory of Prevention and 'Causation' by Omission," *Australasian Journal of Philosophy*, LXXIX, 2 (2001): 216–26.

¹⁵See Helen Beebe, "Causing and Nothingness," in John Collins, Ned Hall, and L. A. Paul, eds., *Causation and Counterfactuals* (Cambridge, MA: MIT Press, 2004), pp. 291–308; and Achille Varzi, "Omissions and Causal Explanations," in Francesca Castellani and Josef Quitterer, eds., *Agency and Causation in the Human Sciences* (Paderborn: Mentis Verlag, 2007), pp. 155–67.

facts that do not involve omissions as causes, and this means that PTR would have to be interpreted accordingly, in a way that respects those facts.

Finally, some philosophers have suggested that there is more than one concept of causation.¹⁶ If that is the case, then of course we should focus on the concept that is relevant in this context, and we should understand PTR as making reference to that concept, whatever it may be. In sum, it is plausible to believe that *some* suitable version of the principle will be available, regardless of one's specific metaphysical commitments about omissions.¹⁷

So now let us return to Sharks and No Sharks. The way to show that the actual sequence is not the same in both cases is to note the role played by the agent's decision not to attempt a rescue in each case. Recall that the agent is blameworthy for that decision in both scenarios. But does the decision have the same *causal results* in both scenarios? Consider the outcome of the child's death. Does the decision causally result in the child's death in both scenarios?

Whereas it is clear that it does in No Sharks, there is a powerful argument that it does *not* in Sharks. I will reconstruct the argument as a reductio. The assumption for the reductio is as follows:

- (1) In Sharks, the agent's decision not to attempt a rescue causally results in the child's death.

Now add these claims (discussed earlier) as premises:

- (2) PTR
- (3) In Sharks, the agent is blameworthy for his decision not to attempt a rescue.

In addition, note that the agent in Sharks had good reason to believe that his decision not to attempt a rescue would result in the child's death, in the way that was to be expected—that is, via the decision's resulting in a failure to do a series of things: jump into the water, swim toward the child, grab him in his arms, and take him back to the shore. (Note, in particular that, given that he was completely unaware of the sharks' presence, he had no reason to believe that he was unable to do

¹⁶ See Ned Hall, "Two Concepts of Causation," in Collins, Hall, and Paul, eds., *Causation and Counterfactuals*, *op. cit.*, pp. 225–76; and Christopher Hitchcock, "Three Concepts of Causation," *Philosophy Compass*, 11, 3 (2007): 508–16.

¹⁷ There are related questions that I cannot get into here. In particular, one might wonder whether views that assign relevance to facts about other possible worlds are consistent with an actual-sequence model of freedom. I think that the best version of an actual-sequence view is, in fact, one that allows for this. I discuss this in detail in *Causation and Free Will*, *op. cit.*

some of these things.) Moreover, we think that the agent is blameworthy for his decision not to attempt a rescue precisely on those grounds, because of the reasons that he had and what he could foresee. Thus, we can add the further premise:

- (4) In Sharks, the agent is blameworthy for his decision not to attempt a rescue partly because he could foresee that such a decision would likely result in the child's death (in roughly the way that we are assuming it did).

Note that these premises entail the following conclusion:

- (5) In Sharks, the agent is blameworthy for the child's death.

(This is because premises (1), (3), and (4) guarantee that the conditions stated in PTR are met, so according to the principle, the agent's blameworthiness for his decision should carry over to the child's death.) But we have been assuming that (5) is false: although the agent is blameworthy for the child's death when there are no sharks, he is not when there are sharks. If this failed to be true, the scenarios would not constitute a PAP-style case, for then there would be no difference in responsibility between them.

This completes the reductio. We may conclude that assumption (1) was false: the agent's decision does *not* causally result in the child's death in Sharks. PTR played an important role in helping us see this because it helped uncover the fact that, however blameworthy the agent may be for his decision in this case, if we think that his responsibility does not carry over to the outcome, then this must be because we regard the decision as causally disconnected from the outcome.¹⁸

Now, if the agent's decision does not causally result in the child's death in Sharks but it does so in No Sharks, then this means that the actual sequence issuing in the child's death is *different* in each case: it includes the agent's decision in one case but not in the other. This seems to be a highly relevant difference! For it means that the agent is causally linked to the child's death in one case in a way that he is not in the other.¹⁹ Thus, the existence of alternative possibilities is not the

¹⁸Or because we regard it as causally disconnected from the outcome *in the relevant sense*. Again, if there is more than one concept of cause, then the principle helps us zero in on the relevant concept. Similarly, if omissions cannot be causes but, say, only quasi-causes, then the claim should be interpreted as a claim about quasi-causation instead of causation.

¹⁹Even an advocate of the alternative-possibilities model should agree that this is a relevant difference. As explained above, embracing the alternative-possibilities view does not commit one to the idea that the existence of alternative possibilities is *all* that matters to freedom; the composition of actual sequences could also matter, and it is extremely plausible to think that it *does* matter.

only potentially relevant difference between Sharks and No Sharks; there is also an important difference in the actual sequence.

Why, then, did it seem plausible to believe that the actual sequence is the same in both cases? Recall that the reasoning for this was the following: The only difference between the two scenarios concerns the presence or absence of the sharks. But the sharks never intervene when they are present, because the agent never jumps into the water. If the sharks never enter the causal chain, and if they are the only difference between the two scenarios, then the causal chain must be the same in both scenarios. However tempting this reasoning may have seemed at first sight, now we can see that it is flawed. For, as we have seen, the difference in the presence of the sharks can result in a difference in the causal chain, although one that does not consist in the sharks *themselves* being part of the causal chain in one case but not in the other. The difference in the causal chain can instead be that *something else* (for example, the agent's decision) is part of the causal chain or fails to be a part of it, depending on whether there are sharks in the water to prevent a rescue attempt.

On reflection, this should not be too surprising, especially when we are dealing with scenarios involving omissions. Imagine a doctor who decides not to operate on a seriously ill patient. If the patient dies, did the doctor's decision not to operate on him contribute to his death? It depends: if the operation would have saved the patient's life, then it did; but if the operation would not have helped—say, because the patient had a deadly tumor in a part of his brain that was inaccessible by surgical means—then it did not. Whether the decision causally resulted in the patient's death depends on factors that never entered the actual chain of events (given that the doctor decided not to operate on the patient), factors such as the accessibility or inaccessibility of the relevant parts of the brain. Similarly, in Sharks and No Sharks, whether the agent's decision not to attempt a rescue causally resulted in the child's death depends on factors that never entered the actual chain of events (given that the agent decided not to attempt a rescue); these factors include the presence or absence of sharks in the water.

The argument can be easily extended to the other scenarios presented in section I. In all cases, the strategy is to show that in one of the members of the pair (the scenario in which the agent lacks alternative possibilities and is not responsible), the relevant decision by the agent does not cause the relevant outcome; thus, the actual sequence is different from the other scenario in that respect. In Light Bulb 2, my decision not to flip the switch does not cause the light to stay off (but it does in Light Bulb 1). In No Phones, my decision not to call the police does not cause the police to not be alerted (but it

does in Phones). And in *All Roads Lead to Rome*, Ryder's decision to steer the horse in a certain direction does not cause some Romans to be hurt (but it does in *Not All Roads Lead to Rome*).²⁰ All of these cases fall together for the same principled reason. And in all of these cases, the appearance that they worked as PAP-style cases can be explained away in the same manner, by drawing attention to how natural it is to overlook the relevance of certain factors to the composition of causal chains.

Note, finally, that all of this is perfectly in line with the central idea behind the actual-sequence model of freedom, to which I alluded briefly in the introduction. Again, the main idea is that freedom is exclusively a function of the actual sequence of events issuing in the agent's behavior. If a factor is not part of the actual sequence and does not help determine what the actual sequence is in any way, then it is irrelevant to the agent's freedom. This is why, in particular, it seemed inappropriate for agents to try to excuse their behavior by drawing attention to the existence of such factors. This is also why, on this model, the presence of the neuroscientist in a Frankfurt-style case is irrelevant to the agent's freedom: because his presence does not affect the actual sequence in any relevant way, the agent still makes his choice purely on the basis of his own deliberation and reasons. Interestingly, as we have seen, the presence of the sharks is *not* like the presence of the neuroscientist. For although the sharks themselves are never part of the actual sequence, they are still relevant to determining *what the actual sequence of events is*, in that case. In the same way that the inaccessibility of the tumor makes the doctor's decision not to operate irrelevant, and, as a result, the causal sequence resulting in the patient's death does not include the doctor's decision, the presence of the sharks makes the agent's decision not to attempt a rescue irrelevant, and, as a result, the causal sequence resulting in the child's death does not include the agent's decision. So the sharks *are* relevant to the actual sequence, in a way that the neuroscientist is not.²¹

²⁰ As I noted, this is the only example of those reviewed that does not involve omissions. The recent causation literature contains an extensive discussion of cases like *All Roads Lead to Rome*—usually called “switches”; see, for example, Stephen Yablo, “De Facto Dependence,” this JOURNAL, XCIX, 3 (March 2002): 130–48; and Ned Hall, “Structural Equations and Causation,” *Philosophical Studies*, CXXXII, 1 (2007): 109–36. I discuss the role of switches in an actual-sequence view of freedom in “Actuality and Responsibility,” *Mind*, CXX, 480 (2011): 1071–97; and in “Making a Difference in a Deterministic World,” *The Philosophical Review*, CXXII, 2 (2013): 189–214.

²¹ For more on this and what is behind this difference, see my *Causation and Free Will*, *op. cit.*, chapters 2 and 3.

III

I have argued that the main candidates for PAP-style cases that can be found in the literature fail and that they fail for the same principled reason. The reason is that, in assuming that the only potentially relevant difference between the two scenarios in the pair is the existence of alternative possibilities (and, in particular, in assuming that there is no difference in the actual sequence of events), they violate a highly plausible principle concerning how agents can become responsible for results. Given other equally plausible assumptions about the cases, the truth of that principle guarantees that the difference in alternative possibilities is *not* the only difference between the scenarios and thus that the examples fail to be genuine PAP-style examples.

Now, of course, this argument only works because it involves cases of derivative responsibility—that is, cases where the agent's responsibility is inherited from other things. The question naturally arises, then: Couldn't there be PAP-style cases in which what is at stake is not derivative but, more fundamentally, *basic* responsibility? Because the locus of basic responsibility is commonly assumed to be decisions, the question becomes: Could there be PAP-style cases in which what is at stake is the agent's responsibility for a *decision*?

Let us think about what this would involve. A PAP-style case of this kind would have to consist of two scenarios: one in which the agent is intuitively responsible for his decision and has alternative possibilities (is able to make a different choice), and another one in which the agent is intuitively not responsible for his decision and lacks alternative possibilities (is unable to make a different choice). Moreover, the two scenarios would have to be such that the difference in alternative possibilities is the *only* potentially relevant difference between them (the only difference that could explain the difference in responsibility). As we have seen, this means, in particular, that there cannot be a difference in the actual sequence of events leading to the decision; in other words, the causal history of the decision would have to be the same in both cases.

But notice what follows from this. In the second scenario, where the agent could not have made a different choice, the factors by virtue of which he was unable to make a different choice (the inevitability factors, whatever these may be) cannot be part of the causal history of the decision. Otherwise the causal history of the decision would *not* be the same in both cases, given that those factors are not present in the first scenario and thus are not part of the causal history of the decision in that case. So the second scenario would have to be one where the new factors are not the explanatory factors. In other words, the second

scenario would have to be a *Frankfurt-style case*. Of course, it need not involve sneaky neuroscientists. But it should involve some kind of backup process that is never activated, which is a characteristic mark of Frankfurt-style cases.

Now, recall that, at the same time, the second scenario has to be one where the agent is intuitively *not* responsible for his decision. This gives rise to an insurmountable problem: arguably, no scenario can meet all of these conditions simultaneously. If a scenario has the structure of a Frankfurt-style case, we will not tend to see the agent as not responsible for his decision; if anything, our initial reaction is likely to be the opposite. Arguably, this is why Frankfurt's paper has been so influential: because people's intuitions about those scenarios tend to align with Frankfurt's own intuition (we tend to see the agent as responsible for his choice). At any rate, note that all that is needed for the problem to arise is that we lack a clear intuition that the agent is *not* responsible for his choice in those cases. And this seems clearly true.

Hence, there are no successful PAP-style cases involving decisions. Looking back, it is not at all surprising that the best candidates for PAP-style cases that can be found in the literature involve results (and, in particular, omissions) instead of decisions. For, as we have seen, their apparent success is only due to the fact that the real nature of the actual sequence can be elusive in those scenarios. This does not happen in Frankfurt-style cases. In Frankfurt-style cases, the composition of the actual sequence is very clear: the decision is caused by the agent's own reasons, in a way that is not affected by the presence of the neuroscientist or the backup process, and this contributes to the intuition that the agent is responsible for his decision in those cases. The moral that I think we should draw from this is that the better we understand the actual sequence, the easier it is to see the failure of the strategy that appeals to PAP-style cases.

I conclude that there are no successful PAP-style cases. Whatever else we should infer from this will depend on how we see the other aspects of the debate. Notably, it will depend on what we think about the initial plausibility of the actual-sequence model versus that of the alternative-possibilities model. If we are naturally drawn to the actual-sequence view because we initially regard it as very plausible, then we should only abandon it if given good reason to believe otherwise. If successful, PAP-style cases would provide such a reason. Therefore, if we can show that they fail, this is an important victory for the actual-sequence view.

CAROLINA SARTORIO

University of Arizona