Article

Frankfurt-Style Cases and the Explanation Condition for Moral Responsibility: a Reply to Swenson

COVA, Florian

Abstract

Frankfurt-style cases (FSCs) are supposed to constitute counter-examples to the principle of alternate possibilities, for they are cases in which we have the intuition that an agent is morally responsible for his action, even though he could not have done otherwise. In a recent paper, Swenson (2015) rejects this conclusion, on the basis of a comparison between standard FSCs, which typically feature actions, and similar cases involving omissions. Because the absence of alternate possibilities seems to preclude moral responsibility in the cases of omissions, and because there is no morally relevant difference between the cases of actions and omissions, Swenson concludes that agents are not morally responsible in standard FSCs. In the present paper, I argue that Swenson's argument fails because there are at least two very important differences between both types of cases. First, there is a difference about whether agents in such cases actually perform the relevant action: while agents actually perform the relevant action in standard FSCs, they do not in FSCs supposedly involving omissions, for omissions require the [...]

Reference

CrossMark

Frankfurt-Style Cases and the Explanation Condition for Moral Responsibility: a Reply to Swenson

Florian Cova[1]

Abstract Frankfurt-style cases (FSCs) are supposed to constitute counter-examples to the principle of alternate possibilities, for they are cases in which we have the intuition that an agent is morally responsible for his action, even though he could not have done otherwise. In a recent paper, Swenson (2015) rejects this conclusion, on the basis of a comparison between standard FSCs, which typically feature actions, and similar cases involving omissions. Because the absence of alternate possibilities seems to preclude moral responsibility in the cases of omissions, and because there is no morally relevant difference between the cases of actions and omissions, Swenson concludes that agents are not morally responsible in standard FSCs. In the present paper, I argue that Swenson's argument fails because there are at least two very important differences between both types of cases. First, there is a difference about whether agents in such cases actually perform the relevant action: while agents actually perform the relevant action in standard FSCs, they do not in FSCs supposedly involving omissions, for omissions require the possibility to have done otherwise. Second, while the agent's behavior in standard FSCs actually explain that he performed the relevant action, the agent's behavior in FSCs including omission actually fails to explain why the agent did not perform the relevant action. Beyond Swenson's argument, I end up discussing what factors ultimately explain (and justify) our intuitions about FSCs involving omissions.

Keywords Alternate possibilities · Frankfurt cases · Moral responsibility · Omissions

One traditional argument for the incompatibility of determinism with moral responsibility rests on the following principle:

✉ Florian Cova
florian.cova@gmail.com

1   Swiss Center for Affective Sciences, University of Geneva, Campus Biotech, Case Postale 60, CH-1211 Genève 20, Switzerland

(PAP) Persons are morally responsible for what they have done only if they could have done otherwise.

Under a certain interpretation of "could have done otherwise", one can deduce from PAP the incompatibility of moral responsibility with determinism for, under that interpretation, an agent in a deterministic world cannot do otherwise.

Against this argument, certain philosophers have argued that PAP is false. One argument for the falsity of PAP relies on a family of thought-experiments called "Frankfurt-style cases" (FSCs). FSCs are cases in which an agent is morally responsible *even if* he could not have done otherwise. This is achieved by constructing cases in which an agent acts entirely *on his own*, but in which a *counterfactual intervener* would have forced him to act the same way if he had not (Frankfurt 1969).

Here is a typical FSC:

Hero—John is walking along the beach and sees a child drowning in the water. John believes that he could rescue the child without much effort. John decides (without deliberating much) to rescue the child, and he successfully does so.

However, unbeknownst to John, a neuroscientist has implanted a chip in Jones's brain which allows him to control Jones's behavior. Had John seriously considered to refrain from rescuing the child, the neuroscientist would have caused him to immediately experience an irresistible urge to rescue the child.

In this case, due to the presence of the neuroscientist, it was impossible for John *not* to save the child. Still, most people have the intuition that, in this kind of case, the agent is morally responsible for his action (Miller and Feltz 2011; Cova 2014; Cova and Kitano 2014). Thus, these counter-examples seem to constitute a counter-example to PAP.

However, in a recent paper, Swenson (2015) has argued that this intuition is misguided, and that FSCs do not constitute convincing counter-examples to PAP. In this paper, I examine this argument and explain how it fails because it ignores two basic conditions for moral responsibility: the performance and the explanation conditions.

1 Swenson's Argument (the Short Version)

Swenson's argument relies on a comparison between the *Hero* case and the following case:

Sharks—John is walking along the beach and sees a child drowning in the water. John believes that he could rescue the child without much effort. Due to his laziness, he decides not to attempt to rescue the child. The child drowns. Unbeknownst to John, there is a school of sharks hidden beneath the water. If John had attempted to rescue the child, the sharks would have eaten him and his rescue attempt would have been unsuccessful.

According to Swenson, it is obvious (i) that John is not morally responsible for failing to save the child in *Sharks*, but (ii) that there is no morally relevant difference

between *Sharks* and *Hero*. Thus, we should conclude, against a widespread intuition, that (iii) John is not morally responsible for saving the child in *Hero*. Because *Hero* is a typical FSC, then this conclusion cast doubt on our intuitions about FSCs in general.

Swenson's argument follows a familiar template. It is a *transfer* argument, which consists in transferring a given verdict about case A to case B, by arguing that there is no relevant difference between the two cases. This kind of argument is widely used in moral philosophy (Singer 1972), and in the free will debate (Pereboom 1995). So, the question is: is it really appropriate to transfer our verdict about *Sharks* to the case of *Hero*? Is there really no morally significant difference between *Sharks* and *Hero*?

Take now the following case:

Normal Hero—John is walking along the beach and sees a child drowning in the water. John believes that he could rescue the child without much effort. John decides (without deliberating much) to rescue the child, and he successfully does so.

In this case, I would say that it is obvious that John is morally responsible for saving the child. But let us now take the following case:

No Hero—John is walking along the beach and sees a child drowning in the water. John believes that he could rescue the child without much effort. John decides (without deliberating much) to rescue the child. However, before he can reach the child, the child is saved by Mark, a local lifeguard.

Here, I think it is obvious that John is *not* morally responsible for saving the child. Now, take the following argument:

1) John is not morally responsible for saving the child in *No Hero*.
2) There is no morally relevant difference between *Normal Hero* and *No Hero*.
3) Thus, John is not morally responsible for saving the child in *Hero*.

Of course, (i) this argument is preposterous, and no one would take it seriously. However, my claim is that (ii) there is no relevant argumentative difference between this argument and Swenson's. Thus, (iii) Swenson's argument fails. More precisely, my claim is that, whatever the relevant difference is between *Normal Hero* and *No Hero*, it also holds between *Hero* and *Sharks*.

In sections 2 and 3, I examine two different ways of expressing the difference between *Normal Hero* and *No Hero*, and in which my dummy no-difference argument could fail. I will explain why Swenson's argument shares the same problems than this argument.

2 The Performance Condition for Moral Responsibility

2.1 The Performance Condition

So, why does the argument I sketched in the previous version fail? Obviously because it is false that there is no relevant difference between *Normal Hero* and *No Hero*. But what is the relevant difference?

Traditionally, philosophers have identified two important conditions for moral responsibility: the "control condition", and the "epistemic condition". Can the difference be located at the level of the control condition? It does not seem so: John seems to have the same amount of control on his action in both cases. Can it be located at the level of epistemic condition? It does not seem so either: John seems to have epistemic access to the same kind of facts in both cases.

At this point, my reader might (rightly) think that I am playing dumb. The difference between the two cases is more than obvious: while John actually saves the child in *Normal Hero*, he does *not* in *No Hero*. Thus, the reason why John is not morally responsible for saving the child in *No Hero* is simply that he does not save him. Obviously, one cannot be responsible for an action he did not perform.

This might be obvious, but this also shows that, in addition to the control and epistemic conditions for moral responsibility, there is also a more basic requirement. Let us call it the "performance principle":

(Performance) An agent A can be morally responsible for φ-ing if and only if A actually φ-ed.

As we will see, it is sometimes important to remember obvious truths.

2.2 What It Takes to Fail

Let us now go back to Swenson's argument. Clearly, John fulfills the "performance condition" for being morally responsible for saving the child in the *Hero* case. But, is it the case that he fulfills this condition for being morally responsible for failing to save the child in the *Sharks* case? Simply put: does John actually fail to save the child in *Sharks*?

To answer this question, we must first realize that "failing to X" is not the same thing as "not X-ing". Imagine a case in which John is at home, and the child is at school, so that nobody is either walking by the water, or drowning in it. In such a case, it would be correct to say that "John does not save the child". However, it would be very strange to say that "John fails to save the child". Similarly, let us imagine that I promised my neighbor I would water her plants, but in fact did not. It would be correct to say that, in this case, I failed to water my neighbor's plants. However, it would seem strange to say that Marcel, a French philosopher who died a few years ago, failed to water my neighbor's plants. Still, it is true that "Marcel did not water my neighbor's plants". Together, these trivial observations points to a very simple conclusion that it takes more to "fail to X" than merely "not X-ing".

What does it take to fail to X? Here is not the proper place to give a full account of failures. However, I think that one requirement is intuitive: for A to fail to X, A must have had the possibility to X. Thus, I propose to take the following principle as intuitive:

(Opportunity) An agent A fails to X only if A had the opportunity to X at some point.

To some, this principle might not seem intuitive. This is because the ordinary use of the word "fail" allows for cases in which this principle is false. For example, if John

had tried to save the child in *Sharks* but did not succeed (because he was eaten by the sharks), we would indeed say that he failed to save the child. It is because, in this sense, failing to X only means that someone tried to X but did not succeed. However, this is not the use we are interested in here: if we were, it would not even make sense to wonder whether John failed to save the child in *Sharks* (given that it is clear that he did not even try to).

Rather, the sense of "failing" we are interested in here is the one according to which "John failed to save the child" is equivalent to "John omitted to save the child", and that, in the *Sharks* case, would also entail that "John let the child die". However, I think it is inappropriate to say that one "omitted to X" when one could not have done X. And I think it would also be inappropriate to say that one let Y happen, when one could not have prevented Y. Thus, I think that the sense of "failing" that is relevant to the present discussion is one in which one cannot fail to X without having the opportunity to do X (see Cova and Naar 2016 for a similar claim).

Let us now go back to *Sharks*. Did John ever have the opportunity to save the child? Clearly not. As stated in the description of the case, it was absolutely impossible for John to save the child. Now, if (Opportunity) is true, this means that John did not actually fail to save the child. But, if this is true, this means that John does not fulfill (Performance) in *Sharks*, which explains why he is not responsible for failing to save the child (that is because he does not actually fail to save the child). In the end, this means that there is a relevant difference between *Sharks* and *Hero*: John fulfills (Performance) in *Hero*, but not in *Sharks*.

Thus, there is a relevant difference between *Hero* and *Sharks*. While John in *Hero* performs the relevant action (saving the child), he does not perform the relevant action (failing to save the child) in *Sharks*. Given that the two cases differ with respect to a basic requirement for moral responsibility, Swenson's argument fails.

2.3 Swenson's Argument (the Extended Version)

So far, for expository reasons, I have only presented a shortened version of Swenson's argument. In its full version, the argument does not only compare *Hero* to *Sharks*, but also makes use of two intermediary cases:

Penned-in Sharks—Everythin occurs just as in *Sharks* except for the fact that the sharks are penned up. However, unbeknownst to John, there is an evil observer who wishes for the child to drown. If John had jumped into the water, the evil observer would have released the sharks, and as a result, the sharks would still have prevented John from rescuing the child. But the presence of the observer plays no role in the actual sequence of events.

Sloth—In this case, there are no sharks present to prevent a rescue by John. The evil observer is now monitoring John's thoughts instead. John decides (without deliberating much) to refrain from saving the child. If John had seriously considered attempting to rescue the child, the evil observer would have caused him to experience an irresistible urge to refrain from saving the child. However, this observer still plays no role in causing John's decision to refrain from attempting a rescue.

Swenson's full argument is then the following:

- (P1) In *Sharks*, John is not responsible for failing to save the child.
- (P2) If John is not responsible for failing to save the child in *Sharks*, then he is not responsible for failing to save the child in *Penned-in Sharks*.
- (P3) If John is not responsible for failing to save the child in *Penned-in Sharks*, then he is not responsible for failing to save the child in *Sloth*.
- (P4) If John is not responsible for failing to save the child in *Sloth*, then he is not responsible for saving the child in *Hero*. Thus,
- (Conclusion) John is not responsible for saving the child in *Hero*.

However, the presence of intermediary cases does not change much to Swenson's argument. Since we have already identified the relevant difference between *Sharks* and *Hero*, we only need to determine when exactly this difference breaks the series of equivalences proposed by Swenson.

Let us begin by the simplest case: the rupture does not occur between *Sharks* and *Penned-in Sharks*. Indeed, it also seems that John does not fail to save the child in *Penned-in Sharks*. Here, as in *Sharks*, John never has the opportunity to save the child, which means that John does not save the child, but does not actually *fail* to save the child. Thus, (Performance) is not fulfilled either in *Penned-in Sharks*.

Thus, as claimed by Swenson, there does not seem to be a difference between *Sharks* and *Penned-in Sharks*: in both cases, John is not morally responsible for the relevant action (failing to save the child) because John does not actually perform this action. This is in line with Swenson's intuition that John is not responsible for failing to save the child in *Penned-in Sharks*.

This means that the relevant difference occurs either between *Penned-in Sharks* and *Sloth*, or between *Sloth* and *Hero*. This is where things get tricky, as determining which option is the right one requires determining whether John actually fails to save the child in *Sloth*. However, whatever the answer to this question, Swenson's argument is doomed. Either John actually fails to save the child and fulfills (Performance) in *Sloth*, and there is a significant difference between *Penned-in Sharks* and *Sloth*, or John does not fail to save the child in *Sloth*, and there is a significant difference between *Sloth* and *Hero*. Either way, Swenson's chain of equivalences is broken.

However, beyond Swenson's argument, it might prove philosophically interesting to determine whether John fails to save the child in *Sloth*. Answering this question depends on how we should interpret and apply (Opportunity). As we saw, John can only be said to have failed to save the child if he had the opportunity, and hence, the possibility to save the child. Did John in *Sloth* had the possibility to save the child?

In one sense, he had not. Indeed, given that it was impossible for John to form the intention to save the child, it was also impossible for John to save the child. However, in another sense, he had. Indeed, even though John could not form the intention to save the child, it is still true that, had John formed the intention to save the child, he would have had a chance to save the child. Determining whether it was possible for John to save the child thus depends on what set of possible worlds is relevant to assess whether John's not saving the child counts as John failing to save the child: the narrow set of possible worlds in which actual world constraints on John's intentions are kept

constant, or the broader set of possible worlds containing worlds in which the counterfactual intervener would not intervened and John is able to form the intention to save the child.

This is a difficult question. However, it must be kept in mind that, despite its troubling resemblance with this traditional question, it is completely independent of how we should interpret possibilities in the PAP. Indeed, it is possible to accept a broad interpretation of possibilities when it comes to determining what counts as a failure, and still accept a more narrow (i.e., incompatibilist) reading of possibilities when it comes to assessing an agent's moral responsibility in light of the PAP. Here, I am only interested in the kind of possibilities relevant to determine when not doing something counts as failing to do something.

I have, however, no definitive answer to offer, though I would argue for selecting the broad set of possible worlds. Thus, I would say that, in *Sloth*, John actually fails to save the child, because it has the relevant possibility, meaning that there are many accessible possible worlds in which he decides to save the child and actually succeeds.

A first reason for this answer is that it seems to best fit our intuitions about these cases, at least if we take Swenson's intuitions to be a representative (more on that later). Indeed, according to Swenson, we have the intuition that John is responsible for failing to save the child in *Sloth*. This means that we have the intuition that John actually fails to save the child. And because it is ceteris paribus better to choose the answer that best fits our intuitions, this gives us a reason to prefer considering the broad set of possible worlds as relevant.

The second reason comes from Clarke (2011): given that we are pretty sure that omissions do occur, it seems strange to make the existence of omissions depend on the uncertain claim that people's actions are not determined. To quote Clarke (2014, p. 92):

> What would the truth of determinism imply about omitting or refraining? I think we can take as resilient data such facts as that people sometimes omit to send holiday greetings or wear their seat belts and that we sometimes abstain, boycott, or fast. Our assurance that these things are so is not threatened by the possibility of determinism.

If we accept these reasons, then we can draw the line between *Penned-in Sharks* and *Sloth*: while John fulfills (Performance) in *Sloth* and *Hero*, he does not in *Sharks* and *Penned-in Sharks*, and this is why we have different intuitions about these cases.

2.4 Swenson's Answers

In a follow-up to his paper, Swenson (2016) discusses the idea that there might be an asymmetry in requirements for moral responsibility between action and omission. His main target is the following principle:

(INTAB) An agent is responsible for omitting to A only if, had the agent intended to A, he would have been able to A.

Note that Swenson has nothing against the truth of INTAB. What troubles him is the idea of (i) accepting INTAB while (ii) holding that agents in FSCs such as *Hero* are morally responsible for their actions. What is the problem? There are two of them, according to Swenson.

The main problem, according to Swenson, is that accepting both INTAB and the success of FSCs leads us to accept an asymmetry between actions (that do not require the ability to do otherwise) and omissions (that do require such an ability). Such an asymmetry, according to Swenson, seems arbitrary.

How is that relevant to the objection I developed in the previous sections? It turns out that, taken together, (Performance) and (Opportunity) implies (INTAB). And, as I support that agents are morally responsible in FSCs such as *Hero*, this leads me to endorse the asymmetry Swenson finds problematic. However, note that, in the present case, there is nothing arbitrary about (INTAB): it is a natural consequence of (Performance), a seemingly uncontroversial principle, and (Opportunity), a substantial but plausible thesis on the nature of omissions. So, I do not think the charge of arbitrariness would be justified.

However, beyond arbitrariness, Swenson also argues that "we ought to regard a lack of symmetry as a cost". Thus, one might argue, even if my proposal justifies in a non-arbitrary way the asymmetry between action and omission, we should still prefer principles that do not have asymmetrical implications.

This is an interesting argument, but it fails to see that any requirement for moral responsibility will in the end create asymmetrical requirements for moral responsibility. Let us suppose that the PAP is true for both actions and omissions (so that there is no asymmetry between them). And let us now compare the two following cases: one in which John deliberately fails to save a child drowning in a pond (while he could have saved him, because he is a good swimmer), and one in which Mark deliberately fails to save a child from being kidnapped (while he could have saved him, because he is a good fighter). We can see that there are still asymmetries in requirements for moral responsibility: John's responsibility for failing to save the drowning child requires the ability to swim, but not to fight, while Mark's responsibility for failing to save the kidnapped kid requires the ability to fight, but not to swim. Of course, one might argue that these are not genuine asymmetries: though the cases feature different requirements for moral responsibility, all requirements can be deduced from the conjunction of one ultimate principle (the PAP) and the particular properties of each action (by its very nature, John's failing to save the drowning child required that John be able to swim, but not Mark's failing to save the kidnapped kid). But, the same is true for the asymmetry between action and omission that I deduced earlier: it can be deduced from the conjunction of one ultimate principle (the performance principle) and the particular properties of each type of actions (by their very nature, omissions do require that agent be able to do otherwise, while full-blown actions do not). Thus, in the end, an account of moral responsibility that rests on the performance principle is no more asymmetrical than an account that would rest on the PAP.

Finally, Swenson has a final objection against those who would want to accept both INTAB and the success of FSCs. It starts with the following case:

Post Intention Frankfurt Case—Black wishes Jones to cast his vote for presidential candidate A. In order to ensure that Jones does this, he implants a chip in Jones's brain which allows him to control Jones's behavior in the voting booth. (Jones has no idea about any of this.) Black prefers that Jones vote for candidate A on his own. But if Jones forms the intention to vote for anyone other than A, Black will immediately use his chip to cause Jones to vote for candidate A instead. As it turns out, though, Jones votes for candidate A on his own and Black never exerts any causal influence on Jones's behavior.

According to Swenson, in this case, it is intuitive that "if Jones is responsible for voting for A, then he is responsible for refraining from voting for B". Thus, one has to either (i) reject the idea that being morally responsible for omission requires the possibility to do otherwise, or (ii) reject the idea that agents are morally responsible for their actions in FSCs.

This is an interesting dilemma. However, the whole argument presupposes that John actually refrains from voting B in this case, and it is far from clear. If "refraining from voting for B" is the same as "failing to vote for B" (as Swenson seems to assume), then, according to (Opportunity), it is not the case that John refrains from voting for B in this case (as he never had the opportunity to vote for B). If, on the contrary, when Swenson speaks about John refraining of voting for B, he only speaks about John's decision not to vote for B, and the action he takes to make sure he is not voting for B (such as voting for A), then John actually refrains from voting for B in this sense, but it is not clear that this is relevant to moral responsibility for omissions.

Thus, one possible difference between *Normal Hero* and *No Hero*, and between *Sharks* and *Hero* might be just that: that John performs the relevant action in *Normal Hero* and *Hero*, but not in *No Hero* and *Sharks*.

3 The Explanation Condition for Moral Responsibility

3.1 Explanation and Responsibility for Events and States of Affairs

There might be another important difference between *Normal Hero* and *No Hero*, and between *Sharks* and *Hero*. But to understand it, we might take a detour by the distinction between responsibility for actions, and responsibility for outcomes (events or states of affairs).

Indeed, Swenson's whole discussion is framed in terms of moral responsibility for what a person does (saving the child, failing to save the child). But a similar discussion could have focused on responsibility for certain outcomes (the child's death, the child's survival). Indeed, some versions of, or replacements for the PAP have been formulated in terms of outcomes. For example, Van Inwagen (1983) puts forward the following principle:

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

Of course, Frankfurt cases are also supposed to be a counter-example to such outcome-based principles. For example, one might say that, in *Hero*, we have the intuition that John is morally responsible for the child being saved, even though he could not have prevented that from happening. However, an argument similar to Swenson's, but put in terms of outcome, could be used to defeat this conclusion: (i) John is not morally responsible for the child's death in *Sharks*, (ii) there is no morally significant difference between *Sharks* and *Hero*, thus (iii) John is not morally responsible for the child's being saved in *Hero*.

Such an argument would seem to bypass the objection developed in section 2: because no direct reference is made to what John *does*, the performance condition

cannot be invoked to explain why John is not morally responsible for the child's death in *Sharks*. The direct equivalent of the performance condition for outcomes would probably be something like this:

(Occurence) An agent A can be morally responsible for an event (or states of affairs) E if and only if E actually occurred (or obtained at some point).

But there is no reason to deny that the child's death actually occurred in *Sharks*, so (Occurence) cannot not help us.

We thus need another explanation. Let us get back to our simple cases: *Normal Hero* and *No Hero*. Why think that John might be responsible for the child being saved in *Normal Hero*, but not in *No Hero*? One simple answer is the following: while John seems to have something to do with the child being saved in *Normal Hero*, he has strictly nothing to do to the child being saved in *No Hero*. To put it otherwise: he is not responsible for the child being saved in *No Hero* because he has in no way contributed to the child being saved.

Now, how are we to translate in more precise terms the intuition that one can be morally responsible for a given outcome only if one has something to do with the occurrence (or the obtaining) of this outcome? One possible way to spell this intuition could be the following:

(Production) An agent A can be morally responsible for an event (or state of affairs) E if and only if A's behavior is a cause of the fact that E occurs (or obtains).

However, this principle has one important shortcoming: it seems that we can be responsible for state of affairs we did not cause. For example, if I promised my neighbor I would water his plants when he is away, and the plants die because I never water them, it seems appropriate to hold me morally responsible for the fact that the plants died. However, it is controversial whether my behavior actually *caused* the plants' death. This is why I think a more appropriate equivalent to the performance principle would be something like the following "explanation principle":

(Explanation) An agent A can be morally responsible for a state of affairs S if and only if (i) S obtains, and (ii) A's behavior (partly or totally) explains the fact that S obtains.

Indeed, even if it is controversial whether my not watering the plants *caused* their death, I think it is more natural to think that it *explains* their death. If someone asks my neighbor *why* her plants died, she would be justified to answer: "*because* my neighbor did not water them". Thus, the explanation principle seems the best way to capture the idea that you cannot be responsible for something if you have nothing to do with it. One proof is that the explanation principle might explain why it is correct to hold John responsible for the child being saved in the *Hero* case, but not in the *No Hero* case. In *Normal Hero*, it is clear that John's behavior explains why the child survived. We can see it by thinking about what would have happened had John not tried to save the child: the child would have drowned. On the contrary, whatever John's behavior in *No Hero*, the child would still have been saved.

This principle also seems able to explain the difference between *Hero* and *Sharks*. Had John not tried to save the child in *Hero*, the child would have drowned. Thus, John's behavior in *Hero* seems to explain the child being saved. However, John's behavior does not seem to explain the fact that the child died in *Sharks*. Indeed, due to the presence of the sharks, there was no way John could have prevented the child from dying. Neither is John the cause of the child dying. Thus, John's behavior seems completely irrelevant to the explanation of the child's death, and John fails to fulfill (Explanation) in *Sharks*.

The same can be said for John's responsibility in *Penned-in Sharks*: due to the presence of the counterfactual intervener, John could not have prevented the child's death. Neither is he the cause of the child's death. So, it seems that John's behavior does not explain the child's death by any mean, which explains why John is not morally responsible for it.

Sloth, however, is a trickier case. Does John's behavior in *Sloth* explain the child's death? The answer is yes: if John had decided and tried to save the child (and thus behaved differently), the child would have survived. Of course, given the presence of the counterfactual intervener, it was necessary that John did not decide to save the child. However, such considerations are irrelevant to the assessment of explanation. Indeed, one can say that the presence of water on earth explains the emergence of life, because life would not have appeared if there had not been water. And the fact that the universe's initial conditions together with the laws of nature made the presence of water on earth unavoidable changes nothing to that. Similarly, the fact that it was necessary that John did not decide to save the child does not prevent the fact that, had John decided to save the child, he would have succeeded, and thus that John's behavior explains the child's death. Thus, (Explanation) appears to be satisfied in *Sloth*, which might explain why people like Swenson have the intuition that John is morally responsible for the child's death in *Sloth*.

3.2 Explanation and Responsibility for the Things We Do

So, the explanation principle allows us to undercut a modified version of Swenson's argument that would focus on moral responsibility for outcomes (the child's death, or the child being saved). But can it also be used against Swenson's original argument, which is cashed out in terms of what agents do (failing to save the child, or saving the child)?—maybe.

A first way to do it might be simply to argue that, in the context of the specific cases used by Swenson, being morally responsible for the child's death implies that one is responsible for failing the child, while being morally responsible for the child being saved implies that one is morally responsible for saving him. Such an inference would not work in any context: there might be cases in which I am responsible for the child's death because I actively killed him, and contexts in which I am responsible for the child being saved because I warned the lifeguard. However, in the context of Swenson's cases, it seems that these inferences are warranted. Thus, because (Explanation) gives us good reason not to treat responsibility for outcomes similarly in *Hero* and *Sharks*, it would also provide us good reasons to treat moral responsibility for what one does differently in both cases.

A second solution would rely on the fact that actions are actually events among others, and that, as such, the explanation principle directly applies to them. Thus, one could argue that John is morally responsible for his saving the child in *Hero* because this event is caused (and thus explained) by his trying to save the child, and his intent to. And one could argue that John is not morally responsible for his not saving the child in *Sharks* because, no matter what he would have done, he could not have prevented his not saving the child, showing that he has nothing to do with his not saving the child.

Thus, in both cases, the explanation principle seems to warrant the conclusion that the agent's moral responsibility about actions should be treated differently in *Hero* and *Sharks*, which gives one more reason to reject Swenson's argument.

Still, before leaving, I would just like to strengthen my case a bit further. So far, I have presented the explanation principle only as a very intuitive principle, and did not try to defend it on other bases. However, it is good to note that there are full-blown theoretical accounts of moral responsibility that integrates similar principles. One such approach is Björnsson and Persson's Explanation Hypothesis, which aims to provide an account of ordinary moral judgments about moral responsibility (Björnsson and Persson 2012; Björnsson and Persson 2013). The complete formulation of their account is the following:

Explanation hypothesis: people take P to be morally responsible for E to the extent that they take E to be an event of type T and take P to have a motivational structure S of type M such that GET, RR, and ER hold:

- *General explanatory tendency (GET):* motivational structures of type M are significant parts of a reasonably common sort of explanation of events of type T.
- *Reactive response-ability (RR):* motivational structures of type M tend to result from or respond in the right way to agents being held responsible for realizing or not preventing events of type T.
- *Explanatory responsibility (ER):* the case in question instantiates the right sort of general explanatory tendency: S is part of a significant explanation of E of the sort mentioned in GET.

As one can see, one can consider the *ER* condition as a more subtle and precise version of (Explanation). Applying it to Swenson's cases would probably yield the same conclusion that John can be morally responsible in *Hero*, because his motivations explain his saving the child (they cause it), but cannot in *Sharks*, because his motivations does not explain his not saving the child (they do not cause it, and his not saving the child does not seem to depend on his motivations in any significant way).

4 What Drives Our Intuitions about Frankfurt-Style Omission Cases?

My main aim in this paper was to show that there are possible morally significant differences between *Hero* and *Sharks*, and thus to undercut Swenson's argument against the reliability of our intuitions in FSCs. I think (or hope) that I have achieved this goal in sections 2 and 3. Those who are only interested in this question can directly skip the present section and directly go to the conclusion.

However, another interesting question remains: which of these morally significant differences actually drive (and explain) our intuitions about FSCs? Swenson admits that he has the intuition that the agent is morally responsible for what he does, though he could not have done otherwise, in cases such as *Sloth* or *Hero*. Though he ultimately dismisses such intuitions as fundamentally misguided, he never explains why we have them. This is one of his argument's other shortcomings: he never presents a convincing error theory for intuitions that seem stable, firm, and deeply shared.

The two principles I have presented in this paper, (Performance) and (Explanation) are not only principles that give us reasons to treat cases such as *Sharks* and *Hero* differently: they might also be explanation for why we treat them differently. To determine whether this is the case, I decided to dabble in experimental philosophy by having people read vignettes similar to the ones used by Swenson, in the hope of determining what kinds of concerns and principles actually drive people's judgments about such cases.

4.1 Materials

Two hundred ninety-five US residents (158 women, 137 men; age mean = 37.6, sd mean = 13.24) were recruited online through Amazon Mechanical Turk and paid $0.4 for their participation. Each participant was randomly assigned to one vignette: *Normal omission*, *Sharks*, *Penned-in Sharks*, *Sloth*, *Hero*, and *Normal Hero* (the texts for all vignettes are presented in Appendix A.1). After reading the vignette, participants were invited to rate their agreement with eight statements (presented in a random order), on a scale ranging from 1 = "Fully disagree" to 7 = "Fully agree", with 4 being marked as "neither agree nor disagree". Statements varied depending on whether the vignette described actions (saving) or omissions (not saving):

- Trying: John deserves blame [praise] for not trying [for trying] to save the child.
- Blame/praise: John deserves blame [praise] for not failing to save [for saving] the child.
- Responsibility: John is morally responsible for the child's death [the child's survival].
- Performance: John omitted to save the child [saved the child].
- Otherwise: John could have done otherwise.
- Ability: John could have saved the child [could have let the child drown].
- Because: the child drowned [was saved] because of John.
- Explanation: John's behavior explains why the child drowned [why the child survived].

4.2 Results

Detailed results are presented in Table 1. Mean answers for particular cases were not my main concern: whether people think the agent is morally responsible for his action in a particular case is of no direct relevance—after all, participants can misinterpret the vignette, fail to grasp important feature, and it is always possible to discuss their judgments on this basis. Rather, my focus was on what explained participants' answers,

Table 1 Mean (and standard deviation) for each question for each vignette. Last line indicates the number of participants for each condition

	Omission	Sharks	Penned-in sharks	Sloth	Hero	Normal hero
Trying	6.17 (1.22)	5.03 (1.81)	5.80 (1.58)	5.50 (1.92)	6.11 (1.38)	6.09 (1.38)
Blame/praise	5.85 (1.53)	4.32 (2.13)	5.35 (1.79)	5.23 (1.95)	6.36 (1.14)	5.98 (1.46)
Responsibility	5.67 (1.50)	4.34 (2.00)	5.24 (1.81)	4.75 (2.13)	5.80 (1.35)	5.53 (1.53)
Performance	6.15 (1.43)	5.5 (1.78)	5.84 (1.57)	5.64 (1.67)	6.51 (0.94)	6.47 (0.95)
Otherwise	6.54 (1.04)	5.97 (1.26)	5.98 (1.42)	4.77 (2.34)	4.55 (2.04)	5.92 (1.53)
Ability	6.37 (1.14)	3.40 (2.16)	4.45 (2.24)	5.02 (2.23)	4.45 (2.14)	5.89 (1.48)
Because	4.89 (1.90)	4.08 (1.96)	4.49 (2.25)	4.25 (1.89)	6.42 (1.01)	6.55 (0.77)
Explanation	4.93 (1.93)	4.21 (2.00)	4.61 (1.94)	4.27 (2.12)	6.27 (1.06)	6.26 (1.08)
N	54	38	51	44	55	54

that is, the factors that best predicted (and possibly explained) their attribution of moral responsibility.

I had two measures of moral responsibility: *blame/praise* and *responsibility*. Because there was a tight link between participants' answers to the two questions (Cronbach's alpha = 0.82), I averaged them in a single score called "responsibility judgments".

Among the factors possibly driving participants' answers were three prime suspects. First, there was the ability to do otherwise, measured through *otherwise* and *ability*, which I averaged in a single *alternate possibility* score (Cronbach's alpha = 0.75). Second, there was the performance condition, measured through *performance*. Third, there was the explanation principle, measured through *because* and *explanation*, which I also averaged into a single score (Cronbach's alpha = 0.88).

To determine which of these factors actually drove participants' judgments about moral responsibility, I ran a multiple linear regression analysis with all three scores (*alternate possibilities*, *performance*, and *explanation*) as predictor variables and participants' responsibility judgments scores as the dependent variable. Results are presented in Table 2.

As indicated by the results of the multiple regression analysis, all three factors (*alternate possibilities*, *performance*, and *explanation*) predicted participants' responsibility judgments. However, all factors did not play an equal role: *explanation* scores were, from far, the best predictor of responsibility judgments, while *alternate possibilities* scores were the weakest.

Table 2 Results of the multiple linear regression analyses. Beta coefficients (β) indicate the strength of the relationship between the given variable and the target variable (responsibility judgments). R^2 = .54

Model	B	SE-B	β	p
Constant	.53	.31	–	.09
Alt. possibilities	.13	.04	.14	<.001
Performance	.31	.05	.27	<.001
Explanation	.45	.04	.51	<.001

Thus, not only do the performance and explanation principles give us reasons to treat *Sharks* and *Hero* differently, but it also turns out that they actually explain why we tend to treat these cases differently.

4.3 Discussion

As stated earlier, participants' mean responsibility judgments in these cases are not very informative: as can be seen from Table 1, a substantial numbers of participants seem not to fully understand the vignettes, and attribute agents the ability to do otherwise, even in cases involving an evil or benevolent counterfactual intervener. This is why I focused on regression analysis and the factors that drove participants' judgments. However, the overall pattern of participant's judgments can be informative, if we focus on differences between cases, and not on the absolute numbers.

If we focus on such patterns, we can see the following: responsibility judgments are at their highest for normal cases that are not FSCs, such as *Omission* and *Normal Hero*. This is perfectly normal. Responsibility judgments are also within the same range for *Hero*, even if participants are much less likely to state the agent could have done otherwise in this case, making it a successful FSC. So far, participants' answers are in line with the intuitions one would expect from the literature (and Swenson's intuitions).

Also in line with such predictions, participants' responsibility judgments were at their lowest for *Sharks*. However, curiously, there was a cutoff between *Sharks* and *Penned-in Sharks*. Indeed, people's judgments about *Penned-in Sharks* and *Sloth* were somewhat beyond the lowest (*Sharks*) and highest (*Hero*, *Normal Hero*, *Omission*). Such a difference goes against Swenson's proposed verdict and the considerations I developed in the previous section. However, it is interesting that the idea that the agent is more morally responsible in *Penned-In Sharks* than in *Sharks* has actually been supported in the literature: Fischer and Ravizza (1998) actually argued that the agent was morally responsible for failing to save the child in *Penned-in Sharks*, but not in *Sharks*. The in-between status of *Penned-in Sharks* in our results seems to parallel philosophers' disagreement on this particular case.

But if participants' intuitions are driven (mostly) by the explanation principle, how can we account for these results? After all, it seems that John's behavior does not explain his not saving the child in *Penned-in Sharks*. But are we really sure of that? Earlier, I reached this conclusion by applying a counterfactual test: I concluded that John's behavior did not explain his not saving the child because John would not have saved the child, whatever his behavior. However, if we follow the same reasoning, we should reach the same conclusion for the evil observer: after all, whatever the observer's behavior (releasing or not the sharks), John would not have saved the child, given that John had decided not to save the child. Thus, by applying the test, we should conclude that neither John nor the observer is responsible for the child's death and John not saving the child, which seems a bit paradoxical.

One explanation for this failure of the counterfactual test in *Penned-in Sharks* might be that this case is a case of "explanatory overdetermination". Explanatory overdetermination occurs when an event is explained (or caused) by multiple factors, each of which is sufficient for explaining its occurrence. In such cases, the counterfactual test typically fails. For example, if we imagine that someone is killed by two bullets that simultaneously land into his heart, we have a case of explanatory

overdetermination: the impact of each bullet is sufficient to explain this person's death. However, in this case, the counterfactual test fails: by applying it, we should conclude that neither the first nor the second bullet explains the person's death.

Cases of explanatory overdetermination are often confusing: some of them are well-known philosophical puzzles. If participants' responsibility judgments are driven by the explanation principle, and *Penned-in Sharks* is such a case, in which it is hard to determine whether John's behavior actually explains his not saving the child, then it is no wonder that people's judgments in this case somewhat in-between. The same can be said for *Sloth*.

But, if this explanation is right, how can we explain the difference people make between *Sharks* and *Penned-in Sharks*? After all, the same can be said for *Sharks*: by applying the counterfactual test, we should reach the paradoxical conclusion that neither John's behavior nor the presence of the sharks in the sea actually explains John's not saving the child. Why do participants seem more certain that John's behavior is not an explanatory factor in *Sharks* than in *Penned-in Sharks*?

This question actually parallels philosophical objections that have been raised against Fischer and Ravizza's proposal of treating *Sharks* and *Penned-in Sharks* differently. In defense of this proposal, Fischer and Ravizza actually argue that, when applying a certain kind of counterfactual test for these cases, the presence of the sharks should be held constant in *Sharks*, but the presence of the counterfactual intervener should be ignored in *Penned-in Sharks*. Of course, this distinction seems somewhat arbitrary. Still, some of our participants seem to follow it. How can this be explained?

One plausible explanation lies in the fact that, when judging such situations, we typically do not envision all possible counterfactuals, but focus on a certain range. Previous empirical studies have shown that, when reflecting on counterfactuals to determine causation and explanation, people focus on factor that are *abnormal*, that is: rule-breaking. Such rules can be statistical (in which case "abnormal" means "unusual") or moral rules (in which case "abnormal" means "immoral") (Knobe and Fraser 2008; Hitchcock and Knobe 2009). Now, the presence and intentions of the counterfactual intervener seems to infringe both statistical (his presence is not usual) and moral rules (he clearly is here to ensure the child's death). This cannot be said for the presence of the sharks in *Sharks*: in this case, the presence of the sharks is more likely to be treated as a background condition, which should be kept constant. It would be interesting to see if people's intuitions about *Sharks* actually change if the case is tweaked to make the presence of the sharks unusual (e.g., they have just been brought here by a tornado).

Of course, one might say that, though explanatorily relevant, such quirks cannot *justify* the fact that we have different intuitions about *Sharks* and *Penned-in Sharks*. I would agree. However, I think it would be a stretch to conclude that the agent is not responsible in *Penned-in Sharks* (because he is not in *Sharks*). Indeed, it could be the contrary: the agent could be morally responsible in *Sharks* (because he is in *Penned-in Sharks*). If the considerations developed in this section are correct, then (i) our intuitions about cases such as *Sharks* and *Penned-in Sharks* are driven by our intuitions about what explains a given event, but (ii) such intuitions are usually confused in cases such as *Sharks* and *Penned-in Sharks* that involve both matters of causation by absence and explanatory overdetermination. When I started this paper, I shared Swenson's intuitions and was confident about my intuitions about the agent

not being responsible in *Sharks* and *Penned-in Sharks*. Now, I have lost this confidence: after all, it might seem correct to say, both in *Sharks* and *Penned-in Sharks*, that John did not save the child *because* he did not try to. If so, why not consider him morally responsible for that? Until we can ground our judgments on a full-blown philosophical account of explanation that applies to such cases, we cannot know.

4.4 Reversing the Argument

In his 2016 paper, Swenson considers the possibility that one might reverse his argument: after all, if there is no relevant difference between *Sharks* and *Hero*, and the agent is morally responsible in *Hero*, then why not simply conclude that the agent is morally responsible in *Sharks*?

To such a reversal of his argument, Swenson has two main lines of objection. The first is that our intuition in *Sharks* is stronger than our intuitions in *Hero*, and that might be because "*Sharks* is a more realistic, down to earth scenario than Original Frankfurt Case". The second is that we have theoretical reasons to doubt our judgment in *Hero* (such as Widerker's "W-defense", or the fact that it conflicts with the PAP), but no reason to doubt our judgment in *Sharks*.

About the first claim, we can see that most of our participants do not share this confidence: participants' responsibility judgments for *Sharks* are clearly around the midpoint and are in fact split in a bimodal distribution, centered around 0 and 6. On the contrary, participants' responsibility judgments for *Hero* are the highest in the study, have a unimodal distribution, and are centered around 6. Thus, intuitions about *Hero* seem strongest and more stable and widespread than intuitions about *Sharks*. This might be due to the fact that, when comparing *Sharks* and *Hero*, Swenson does not focus on the good form of "realism". Sure, Frankfurt cases involve some cases of sci-fi elements, such as mindreading and mind-controlling devices that do not exist in the real life. However, we often see and reflect about this kind of device in sci-fi movies and literature, and their widespread presence shows that such situations are quite understandable. We seem to have no trouble to judge that the character who is mind-controlled by an evil character is not responsible for his action. No, if FSCs might seem puzzling and confusing, it has nothing to do with the sci-fi stuff: it has to do with things like explanatory overdetermination (the fact that the agent would have been forced to do it anyway, etc.) and counterfactual considerations (contrary to typically evil sci-fi characters, counterfactual interveners in FSCs do not act). However, I contend that such problems are not only also present in cases such as *Penned-in Sharks* and, to a lesser extent, *Sharks* but they are also made worse. Indeed, in typical FSCs, despite all the apparent troubles, it is still clear that the agent's decision actually causes the action, which helps us determine easily whether the agent plays a part in the explanation of his action. In case such as *Sharks* and *Penned-in Sharks*, there are no such things, which may contribute to make them even more alien.

As for the theoretical reasons against our intuition in *Hero* and typical FSCs, there sure are some. However, the fact that there are no theoretical reasons against our intuitions about *Sharks* might only be due to the fact that they have not been scrutinized enough (after all, they do not pose an immediate threat to the PAP). In this section, I have given theoretical reasons to doubt them, namely, that they

might rely on further intuitions about what explains what, and that, barring an appropriate philosophical account of explanation, we have no reason to completely trust these intuitions. So, in its current state, and even if we acknowledged there was no significant difference between *Hero* and *Sharks* (which, as we saw, is not the case), it still would not be clear what Swenson argument actually proves that the agent is not morally responsible in *Hero*, or that he is morally responsible in *Sharks*.

5 Conclusion

Swenson's argument can be summed up in the following way: while responsibility in FSCs such as *Hero* does not seem to require the presence of alternate possibilities, responsibility for omissions in the case such as *Sharks* seems to require the presence of alternate possibilities. Because there is no morally relevant difference between cases such as *Hero* and cases such as *Sharks*, this apparent inconsistency sheds doubt on our intuitions about FSCs such as *Hero*.

Swenson is not the first to note the difference between cases involving actions such as *Hero* and cases involving omissions such as *Sharks*. Several philosophers have interpreted this contrast as showing that there is an asymmetry between action and omission: while moral responsibility for actions does not require alternate possibilities, responsibility for omissions does (Clarke 2011; see Sartorio 2005 for an alternate account of the asymmetry).

In this paper, I have taken a different approach to this so-called "asymmetry". I have tried to show that this asymmetry can be explained by a single, intuitive principle: either the performance principle, according to which agents can only be morally responsible for actions they actually perform, or the explanation principle, according to which agents can only be morally responsible for events or states of affairs that are explained by their behavior. According to both approach, moral responsibility for both actions and omissions have the same requirements: the epistemic condition, the control condition, and either the performance or explanation condition. This means that there is no real asymmetry between actions and omissions: they only have different ways of fulfilling these different requirements.

Thus, against accounts that claim that requirements for moral responsibility should be different for actions and omissions, I agree with Swenson to think that cases of actions and omissions should be treated equally. However, I disagree with Swenson when he claims that our intuitions about moral responsibility in cases of actions and omissions are incoherent: they are perfectly coherent, once taken into account (i) what it means for omission to be performed, or (ii) what it takes for an absence of action to explain something else.

Going further, I have also shown that the fact that we treat cases like *Sharks* and cases like *Hero* differently cannot only been justified, but also explained by the performance and explanation principle, the latter being the main determinant of our intuitions about such cases. Overall, I consider that these arguments and data offer a compelling argument against the claim that there is no morally relevant difference between cases like *Sharks* and cases like *Hero*, thus undercutting Swenson's argument against the validity of our intuitions about FSCs.

Appendix

Cases used for section 4's study

Omission—John is walking along the beach and sees a child drowning in the water. John believes that he could rescue the child without much effort.

Due to his laziness, John decides (without deliberating much) not to attempt to rescue the child. As a result, the child dies from drowning.

Sharks—John is walking along the beach and sees a child drowning in the water. John believes that he could rescue the child without much effort.

Due to his laziness, John decides (without deliberating much) not to attempt to rescue the child. As a result, the child dies from drowning.

Unbeknownst to John, there is a school of sharks hidden beneath the water. If John had attempted to rescue the child, the sharks would have eaten him and his rescue attempt would have been unsuccessful.

Penned-in Sharks—John is walking along the beach and sees a child drowning in the water. John believes that he could rescue the child without much effort.

Due to his laziness, John decides (without deliberating much) not to attempt to rescue the child. As a result, the child dies from drowning.

Unbeknownst to John, there is an evil observer who wishes for the child to drown. If John had jumped into the water, the evil observer would have released a school of sharks, and as a result, the sharks would have prevented John from rescuing the child. However, since John has not even tried to save the child, he did not have to release the sharks.

Sloth—John is walking along the beach and sees a child drowning in the water. John believes that he could rescue the child without much effort.

Due to his laziness, John decides (without deliberating much) not to attempt to rescue the child. As a result, the child dies from drowning.

Unbeknownst to John, there is an evil observer who wishes for the child to drown and is monitoring John's thoughts, thanks to a machine he previously implanted in John's brain. If John had seriously considered attempting to rescue the child, the evil observer would have caused him to experience an irresistible urge to refrain from saving the child. However, since John has not even considered trying to save the child, he did not have to trigger such an urge.

Hero—John is walking along the beach and sees a child drowning in the water. John believes that he could rescue the child without much effort.

John decides (without deliberating much) to rescue the child, and he successfully does so. As a result, the child is saved.

Unbeknownst to John, there is a benevolent observer who wishes for the child to be saved and is monitoring John's thoughts, thanks to a machine he previously implanted in John's brain. If John had seriously considered attempting to rescue the child, the evil observer would have caused him to experience an irresistible urge to rescue the child.

However, since John has not even considered letting the child drown, he did not have to trigger such an urge.

Normal Hero—John is walking along the beach and sees a child drowning in the water. John believes that he could rescue the child without much effort.

John decides (without deliberating much) to rescue the child, and he successfully does so. As a result, the child is saved.

References

Björnsson, G., & Persson, K. (2012). The explanatory component of moral responsibility. *Noûs, 46*(2), 326–354.

Björnsson, G., & Persson, K. (2013). A unified empirical account of responsibility judgments. *Philosophy and Phenomenological Research, 87*(3), 611–639.

Clarke, R. (2011). Omissions, responsibility, and symmetry. *Philosophy and Phenomenological Research, 82*(3), 594–623.

Clarke, R. (2014). *Omissions: agency, metaphysics, and responsibility.* Oxford University Press.

Cova, F. (2014). Frankfurt-style cases user manual: why Frankfurt-style enabling cases do not necessitate tech support. *Ethical Theory and Moral Practice, 17*(3), 505–521.

Cova, F., & Kitano, Y. (2014). Experimental philosophy and the compatibility of free will and determinism: A survey. *Annals of the Japan Association for Philosophy of Science, 22*, 17–37.

Cova, F., & Naar, H. (2016). Do intuitions about Frankfurt-style cases rest on an internalist prejudice?. *Philosophical Explorations, 19*(3), 290–305.

Fischer, J. M., & Ravizza, M. (1998). *Responsibility and control: a theory of moral responsibility.* Cambridge: CUP.

Frankfurt, H. (1969). Alternate possibilities and moral responsibility. *Journal of Philosophy, 66*, 829–839.

Hitchcock, C., & Knobe, J. (2009). Cause and norm. *Journal of Philosophy, 106*(11), 587–612.

Knobe, J. & Fraser, B. (2008) Causal judgment and moral judgment: Two experiments. In: *Moral psychology, vol. 2: The cognitive science of morality: Intuition and diversity*, ed. W. Sinnott-Armstrong, pp. 441–8. MIT Press.

Miller, J. S., & Feltz, A. (2011). Frankfurt and the folk: an experimental investigation of Frankfurt-style cases. *Consciousness and Cognition, 20*(2), 401–414.

Pereboom, D. (1995). Determinism al dente. *Noûs, 29*(1), 21–45.

Sartorio, C. (2005). A new asymmetry between actions and omissions. *Noûs, 39*(3), 460–482.

Singer, P. (1972). Famine, affluence, and morality. *Philosophy & Public Affairs, 1*(3), 229–243.

Swenson, P. (2015). A challenge for Frankfurt-style compatibilists. *Philosophical Studies, 172*(5), 1279–1285.

Swenson, P. (2016). The frankfurt cases and responsibility for omissions. *The Philosophical Quarterly, 66*, 579–595.

van Inwagen, P. (1983). *An essay on free will.* Oxford: Clarendon.