The Philosophical Quarterly Vol. 66, No. 264 *2016*
ISSN 0031-8094
doi: 10.1093/pq/pqv127
Advance Access Publication 11th January 2016

Downloaded from https://academic.oup.com/pq/article/66/264/579/2461022 by William & Mary Libraries user on 04 October 2021

THE FRANKFURT CASES AND RESPONSIBILITY FOR OMISSIONS

By Philip Swenson

In previous work, I presented a challenge for philosophers who appeal to the Frankfurt-style cases (FSCs) in order to undermine the Principle of Alternative Possibilities (PAP). My challenge relied on the claim that there are cases of omitting to act in which the agent is not responsible for her behavior (or lack thereof) and which should yield the same verdict regarding responsibility as the Frankfurt-style cases. In this paper I take a closer look at particular accounts of responsibility for omissions on offer in the literature and argue that they fail to overcome my challenge. In particular I focus on accounts offered by Fischer and Ravizza, Randolph Clarke, and Carolina Sartorio.

Keywords: Frankfurt Cases, moral responsibility, PAP.

I. THE CHALLENGE

Here is a case quite similar to the one originally presented by Frankfurt (1969):

> **Original 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 starts to become inclined 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.

Initially, it would seem that both of the following are true. (1) Jones is morally responsible for voting for candidate A. And (2) Jones could not have done otherwise than he in fact did. Thus, we have an apparent counterexample to PAP. My challenge relies on cases such as the following.

> **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 Author 2016. Published by Oxford University Press on behalf of The Scots Philosophical Association and the University of St Andrews. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com

the child, the sharks would have eaten him, and his rescue attempt would have been unsuccessful.[1]

In Sharks, it seems clear that John is not morally responsible for *failing to save the child*. But if no sharks were present John would have been responsible for *failing to save the child*. So the counterfactual intervention of the sharks does seem to affect John's responsibility. The challenge for defenders of FSCs is to account for both Black's lack of impact on responsibility in Original Frankfurt Case and the sharks impact on responsibility in Sharks. To bring out the forcefulness of this problem, we should examine a string of cases which gradually bridge the gap between Sharks and Original Frankfurt Case. Here are the cases:

Penned-in Sharks: Everything 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.[2]

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.[3]

and:

Hero: John decides (without deliberating much) to rescue the child, and he successfully does so. Unbeknownst to him, if he had seriously considered refraining from rescuing the child, our now benevolent observer would have caused him to immediately experience an irresistible urge to rescue the child.[4]

Hero is structurally identical to Original Frankfurt Case. So defenders of FSCs must claim that in one of these cases John ceases to be responsible. We can make the challenge more precise by looking at the **No Principled Difference Argument**:

(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.

[1] This case is drawn from Fischer & Ravizza (1998: 125).
[2] This case is also from Fischer & Ravizza (1998: 138). They credit David Kaplan for suggesting the case.
[3] This sort of case was suggested by Frankfurt (1994).
[4] This case is drawn from Fischer & Ravizza (1991)

Downloaded from https://academic.oup.com/pq/article/66/264/579/2461022 by William & Mary Libraries user on 04 October 2021

(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.

To account for the purported difference between Sharks and the FSCs either (P2), (P3) or (P4) would have to be rejected. I have previously identified, in Swenson (2015), a general reason for thinking that defenders of FSCs cannot plausibly reject any of these premises.

Following Frankfurt (1969) and Fischer (2010), I have suggested that the reason Black appears to be irrelevant in Original Frankfurt Case is precisely because he does not make anything happen in the actual sequence of events. Fischer (2010) has pointed us to an important distinction between the 'A-Factors' of a situation, which bring about a particular event, and the 'B-Factors' which render the event inevitable but need not cause or bring about the event. Using this terminology, I suggest that the principle underlying our intuitive reaction to FSCs is that *mere B-Factors are irrelevant to moral responsibility*.

The *General Problem* I identified for those who wish to defend FSCs and reject either (P2), (P3) or (P4) is that each case appealed to in the No Principled Difference argument centrally involves the presence of a mere B-Factor (the sharks in Sharks, the evil observer in Penned-in Sharks, etc.). Furthermore, accepting that the agent is not responsible in any of these cases apparently involves rejecting the claim that mere B-Factors are always irrelevant to moral responsibility. Since defenders of FSCs should say that the principle underlying our intuitive reaction to FSCs is correct, it will apparently be difficult for them to accept the claim that John is not responsible in any of the cases appealed to in the No Principled Difference argument. Swenson (2015) provides a more detailed discussion of this *General Problem*.

In the next three sections, I will examine accounts of responsibility for omissions on offer in the literature which would underwrite rejecting each of (P2), (P3) and (P4). I will argue that, in addition to running afoul of the *General Problem*, each account faces significant difficulties of its own. In the final section, I will consider the more radical possibility of rejecting (P1) and offer reasons against doing so.

II. FISCHER AND RAVIZZA AND (P2)

Fischer & Ravizza (1998) have presented a detailed account of responsibility which purports to provide a motivation for rejecting (P2). Appealing to

Downloaded from https://academic.oup.com/pq/article/66/264/579/2461022 by William & Mary Libraries user on 04 October 2021

FSCs, they suggest that, just as we must (on their view) 'hold fixed' the non-intervention of Black in evaluating Jones's responsibility for his action, so too must we hold fixed the non-occurrence of some events in evaluating an agent's responsibility for an outcome. To answer the question of which events must be held fixed, they introduce the notion of a triggering event. A triggering event '(relevant to some consequence C) [is an] event which is such that, if it were to occur, it would *initiate* a causal sequence leading to C'.[5] For example, Black's use of his device to cause Jones to vote for A would count as a triggering event relative to the consequence that a vote is cast for A. Fischer and Ravizza then provide us with the following necessary condition for an agent's being responsible for the consequences of an omission, such as John's failing to save the child:

> Suppose that in the actual world an agent *S* moves his body in way *B* at time *T* via a type of mechanism *M*, and *S's* moving his body in way *B* at time *T* causes some consequence-universal *C* to obtain at *T*+*i* via a type of process *P*. . . [Then *S* is only responsible for *C* on the condition that] If S were to move his body in way *B** [which cannot be identical to *B*] at *T*, and all other triggering events (apart from *B**) that do *not actually* occur between *T* and *T*+*i* were *not* to occur, and a *P*-type process were to occur, then C would not occur.[6]

Fischer and Ravizza argue that we can use this requirement to show that John can be responsible for failing to save the child in Penned-in Sharks even though he cannot be responsible for this failure in Sharks. In Penned-in Sharks, the evil observer's release of the sharks from the pen counts as a triggering event, and so (given Fischer and Ravizza's account) we should hold fixed its non-occurrence. Thus, we get the result that (assuming that John meets all other requirements for being morally responsible) John is responsible for the fact that the child drowned. And this would apparently entail (in this context) that John is responsible for failing to save the child.

Fischer and Ravizza want to maintain that their account yields the result that John is not responsible in Sharks. But it is unclear how they can get this result since it would appear that the shark's sensing that John entered the water should count as a triggering event as well. Fischer and Ravizza discuss this worry in a footnote and they say the following:

> . . . in the alternate sequence, John's jumping into the water would antedate and lead to the shark's sensing that he had done so: thus, the shark's sensing John would not "initiate" – in the relevant sense – the sequence leading to the child's not being saved by John (and thus would not be a triggering event).[7]

[5] Fischer & Ravizza (1998: 110–1).

[6] Fischer & Ravizza (1998: 112 (also 135). Note that I have omitted to state parts of Fischer and Ravizza's account which do not concern us here but which are important for evaluating their account in other contexts.

[7] Fischer & Ravizza (1998: 136).

Downloaded from https://academic.oup.com/pq/article/66/264/579/2461022 by William & Mary Libraries user on 04 October 2021

I do not see how this reply will help (at least with the goal of distinguishing between Sharks and Penned-in Sharks in mind). This is because if the sharks' sensing that John jumped into the water does not count as a triggering event because it is antedated and caused by John's act, then surely the evil observer's (in Penned-in Sharks) sensing that John jumped into the water would not count as a triggering event for the same reason. Thus, it does not appear that Fischer and Ravizza's view (in its current form) can account for the purported difference between Sharks and Penned-in Sharks. (Note that this critique of Fischer and Ravizza is not original. Byrd (2007) and Clarke (2014) press very similar objections.)

Byrd (2007) has recognized that Fischer and Ravizza's unmodified account fails to licence a rejection of claims like (P2), and he suggests a modification that would do the trick.[8] Byrd's view is that we should only hold fixed the non-occurrence of triggering events that are the choices of rational agents. Thus, we hold fixed the non-occurrence of our evil observer's decision to let the sharks out of the pen (in Penned-in Sharks). But we do not hold fixed the non-occurrence of the sharks sensing and attacking John (in Sharks).[9] (Byrd apparently has in mind a construal of rationality on which sharks do not count as rational agents.)

I do not think Byrd's suggestion is satisfactory. Byrd claims that this view 'provides a systematic solution which matches one's judgments in the clear cases and gives proper guidance in the tougher ones'.[10] However, I think that there are clear cases in which Byrd's view yields the intuitively wrong results. Consider for example:

> **Non-Agential Sloth**: In this case, there are no sharks or evil observers present to prevent a rescue by John. John is afflicted by a phobia of water of which he is completely unaware and not responsible for possessing. John decides (without deliberating much) to refrain from saving the child. If John had seriously considered attempting to rescue the child, his phobia would have caused him to experience an irresistible urge to refrain from saving the child. However, this phobia plays no role in causing John's decision to refrain from attempting a rescue.[11]

This case seems to me to call for the same verdict as the original Sloth case. I would be very surprised if many philosophers had the intuition that John is responsible in Sloth but not in Non-Agential Sloth. Yet, this is just what Byrd's

[8] Strictly speaking, Byrd's view seeks to rescue Fischer and Ravizza's account from a parallel problem having to do with responsibility for consequences. And his modified notion of triggering events is thus intended to apply only to responsibility for consequences. However, in order to see whether it can provide Fischer and Ravizza with a way to reject (P2), we can consider a modified version that applies to omissions as well.

[9] Byrd (2007).

[10] Byrd (2007: 63).

[11] This non-agential version is closer to how Frankfurt (1994) originally presented the Sloth case.

Downloaded from https://academic.oup.com/pq/article/66/264/579/2461022 by William & Mary Libraries user on 04 October 2021

account appears to suggest, since it would appear that we should hold fixed the non-occurrence of the evil observer's choices but not the non-occurrence of the irresistible urge.[12] Thus, I do not find Byrd's approach to be very promising.

Now, it is of course possible that some other modification of Fischer and Ravizza's account would be able to distinguish between Sharks and Penned-in Sharks. But I see no obvious way of making such a modification. Furthermore, even if this sort of modification were developed, Fischer and Ravizza's account could still be faulted for running afoul of the *General Problem*. Since their account treats the sharks (in Sharks) as relevant to moral responsibility, it does not do justice to the fact that Black is irrelevant *because* he is a mere B-Factor.

III. CLARKE AND (P3)

Having criticized Fischer and Ravizza's rejection of (P2), I will now consider an account that, if accepted, would licence a rejection of (P3).

> (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.

Randolph Clarke has defended a view on which (P3) will turn out to be false. He proposes the following necessary condition for an agent's being responsible for an omission:

> **INTAB** An agent is responsible for omitting to A only if, had the agent intended to A, he would have been able to A.[13]

Clarke maintains that John is not responsible in Sharks and Penned-in Sharks but is responsible in Sloth. **INTAB** rules out John's being responsible in Sharks and Penned-in Sharks, but allows that John might be responsible in Sloth. This is because in Sloth the counterfactual intervention is set to occur before John forms the intention to save the child. Clarke endorses **INTAB** primarily on the basis that it gets the right results in these and similar cases. I agree with Clarke that **INTAB** yields intuitively correct results in a range of omissions cases. However, I think that it is ultimately implausible for two reasons.

First, as Clarke acknowledges, the truth of **INTAB** would appear to leave us with a significant asymmetry between actions and omissions. The problem is that there does not appear to be any requirement similar to **INTAB** that

[12] Swenson has pointed out that his requirement that the triggering event be the choice of a rational agent could be limited to cases in which the triggering event occurs post choice. This version of the requirement would avoid yielding an incorrect verdict in Non-Agential Sloth, but at the cost of apparent arbitrariness. Furthermore, the revised requirement would still be vulnerable to a Non-Agential version of Penned-In Sharks.

[13] Clarke (2011). McIntyre (1994) defends a similar view. Note that Clarke thinks that **INTAB** is true only in an 'appropriately restricted, revised and refined form'. None of these restrictions, etc. will be relevant to my discussion of **INTAB**.

Downloaded from https://academic.oup.com/pq/article/66/264/579/2461022 by William & Mary Libraries user on 04 October 2021

holds true in the case of action. The parallel to **INTAB** in the case of action would be:

> **ACTION INTAB** An agent is responsible for A-ing only if, had the agent intended to refrain from A-ing, he would have been able to refrain from A-ing.

But this principle appears to be undermined by FSCs in which the counterfactual intervener is prepared to intervene immediately after the intention to refrain is formed.[14] Consider:

> **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.

This case appears to show that **ACTION INTAB** is false.[15]

Clarke is sensitive to this issue. He notes that '**INTAB** imposes a requirement concerning ability that, if Frankfurt is right, has no parallel in the case of action'.[16] So accepting **INTAB** results (by Frankfurtian lights) in a significant asymmetry between what is required in order to be responsible for an action and what is required in order to be responsible for an omission.

This leads to my second reason for rejecting Clarke's account. Consider what Clarke's account would say about Post Intention Frankfurt Case. Since we are rejecting **ACTION INTAB** we will say that Jones is responsible for voting for candidate A. Now suppose that in casting his vote for A, Jones thereby refrained from voting for candidate B. It seems clear to me that in this case Jones' responsibility for voting for A and his responsibility for refraining from voting for B should stand or fall together. If Jones is responsible for voting for A, he is thereby responsible for refraining from voting for B. But if we accept **INTAB**, we cannot say that Jones is responsible for refraining from voting for B, since it is not true that if Jones had intended to vote for B, he would have been able to.[17]

Clarke has pointed out to me that Jones' responsibility for voting for A and his responsibility for refraining from voting for B will not stand or fall together; if Jones is unaware that he has the option of voting for B. So, Clarke suggests, it is open to him to say that responsibility for an action and a corresponding

[14] Fischer points this out in 2008.
[15] At least by the lights of those who accept our basic intuitions concerning FSCs.
[16] Clarke (2011: 621).
[17] Thanks to Patrick Todd for helpful discussion on this point.

Downloaded from https://academic.oup.com/pq/article/66/264/579/2461022 by William & Mary Libraries user on 04 October 2021

omission will not stand or fall together for other reasons as well.[18] I agree that in some cases they will not stand or fall together. What I wish to maintain is that is implausible to hold that they do not stand or fall together in cases where the agent meets the following two conditions. (1) The agent knows that by voting for A he is thereby refraining from voting for B. And (2) the agent either has both the ability to vote for B and the ability to refrain from voting for A *or* lacks both the ability to vote for B and the ability to refrain from voting for A. Both of these conditions can be met in a suitably filled out version of Post Intention Frankfurt Case.

So I have two distinct objections to Clarke's account. First, his account leaves us with a significant asymmetry between actions and omissions. Second, as a result of this asymmetry, Clarke's account fails to yield the result that (in Post Intention Frankfurt Case) if Jones is responsible for voting for A, then he is responsible for refraining from voting for B. (I regard the second objection as the more serious of the two.)

Clarke attempts to lessen the implausibility of positing this asymmetry between responsibility for actions and responsibility for omissions by suggesting that omissions may not be events (as actions are), but instead are the absences of certain actions. He then reasons as follows:

> If such a view is correct, it is to be expected that there might well be major differences between what is required for responsibility for actions and what is required for responsibility for omissions, for an omission, unlike an action, is not an agent's exercising control in bringing something about. Responsibility for omissions will not be a special case of responsibility for actions, for omission isn't a special case of action (anymore than the absence of a lion is a special kind of animal or the absence of red is a special kind of color).[19]

While this may help some, since it creates more separation between action and omission, it does not, at least in my view, remove the unsatisfactory arbitrariness of the asymmetry. Even granting Clarke's view of the metaphysics of omissions, we still lack an explanation for why the metaphysical differences between actions and omissions result in different normative facts concerning the control requirement for moral responsibility.

In defence of his account, Clarke notes that it may be appropriate to sacrifice some degree of symmetry in our theory of moral responsibility in order to account for our intuitions about different cases. I think that this is correct. However, I think that we ought to regard a lack of symmetry as a cost, especially when we cannot see why the asymmetry holds. Furthermore, as we have seen, I do not think Clarke's view accounts for all of our intuitions about cases. It does not yield the result that (in Post Intention Frankfurt Case) if Jones

[18] Conversation with Randolph Clarke. Thanks to Clarke for helpful discussion here.
[19] Clarke (2011: 622).

Downloaded from https://academic.oup.com/pq/article/66/264/579/2461022 by William & Mary Libraries user on 04 October 2021

is responsible for voting for A, then he is responsible for refraining from voting for B. Thus, the *accounting for cases* motivation for rejecting symmetry is at least partially undercut.

In his recent (2014) book, Clarke suggests that there may be another revealing asymmetry between Sharks and Sloth. In Sloth, it is clear that John intentionally refrains from saving the child. But Clarke raises the possibility that in Sharks John does not intentionally refrain from saving the child. Clarke suggests that the fact that 'No matter what he had intended, and no matter how hard he might have tried, he wouldn't have been able to carry out the rescue'[20] makes it the case that John does not intentionally omit to rescue the child. Clarke grants that agents can be responsible for omissions done unintentionally. But he suggests that perhaps:

> what explains why John doesn't intentionally not save the child also explains why he isn't responsible for failing to do so do: no matter what John had intended, and no matter how hard he might have tried, he wouldn't have been able to save the child.[21]

I don't think this suggestion helps Clarke's view to avoid the objection I have raised. After all, in Post Intention Frankfurt Case, it is true that 'no matter what Jones had intended, and no matter how hard he might have tried, he wouldn't have been able to refrain from voting for A'. But I take it that Clarke would say that Jones intentionally voted for A and that he is responsible for voting for A. Thus, our puzzling asymmetry remains.

Finally, it is worth observing that Clarke's approach runs afoul of the *General Problem* by yielding the result that some mere B-Factors are relevant to moral responsibility. The presence of the evil observer in Penned-in Sharks is apparently a mere B-Factor, but, given **INTAB**, it is not irrelevant to responsibility. One could claim that **INTAB** provides us with an explanation for why some B-Factors are relevant while others are not. But this appears inadequate. The FSCs drive us to think that the effect various factors have on which counterfactuals are true of us is irrelevant to our moral responsibility when those factors do not contribute causally to the actual sequence of events. If we accept this intuition, then it seems that we should conclude that the presence of the evil observer is likewise irrelevant to moral responsibility.

III.1. Interlude: Sartorio to the rescue?

Carolina Sartorio has defended a set of claims about responsibility which, if true, may underwrite a rejection of (P3), and avoid the objections I have just raised to Clarke's view. Sartorio endorses the following:

[20] (Clarke 2014: 152)
[21] (Clarke 2014: 153)

Downloaded from https://academic.oup.com/pq/article/66/264/579/2461022 by William & Mary Libraries user on 04 October 2021

TR(Causal): an Agent's responsibility for X transmits to an outcome Y iff X causes Y (and the usual provisos obtain).[22]

She also argues for:

> NA(Causal): An action can cause an outcome even if the outcome would still have occurred in the absence of the action. By contrast an omission cannot cause an outcome if the outcome would have still occurred in the absence of the omission.[23]

TR(Causal) and NA(Causal) together entail:

> NA: An agent's responsibility for an action can transmit to an outcome even if the outcome would have occurred anyway in the absence of the action. However, an agent's responsibility for an omission cannot transmit to an outcome if the outcome would have occurred anyway in the absence of the omission.[24]

If NA is correct, then it is plausible that John is not responsible in Penned-in Sharks. John plausibly does not omit to save the child if he tries and fails. So it is true that in the absence of the omission the child would still have drowned. And since I take it that John is responsible for failing to save the child only if he is responsible for the child's drowning, John would not be responsible for failing to save the child.

Furthermore, we could now explain why **INTAB** may be true even though **ACTION INTAB** is false. The asymmetry is explained by the fact that actions can cause outcomes under conditions in which omissions cannot cause outcomes. And perhaps most significantly, we could also now offer a response to the *General Problem,* and a more satisfying explanation for why some B-Factors are relevant to moral responsibility. Perhaps, what rules out Black's relevance in Original Frankfurt Case is that he makes no difference to the causal structure of the actual sequence. A-Factors can be relevant to responsibility because they cause something to occur in the actual sequence. And B-Factors can be relevant because they can prevent omissions from causing events in the actual sequence (thus, unlike Black, altering the causal structure of the actual sequence). The presence of the intervener in Penned-in Sharks does make a difference to what happens in the actual sequence of events because he prevents John's omission from being a cause of the child's drowning. In Sloth, by contrast, John's omission can be a cause of the child's drowning because it is not true that the child would still have drowned in the absence of the omission.[25] Thus, a rejection of (P3) would appear to be vindicated.

[22] Sartorio (2005: 468).
[23] Sartorio (2005: 470).
[24] Sartorio (2005: 470).
[25] This is because in Sloth the evil observer stops John from rescuing the child by getting him to omit to rescue the child. It is somewhat surprising that Sartorio's NA has this consequence since she is sceptical of John's responsibility in Sloth (see Section IV).

Downloaded from https://academic.oup.com/pq/article/66/264/579/2461022 by William & Mary Libraries user on 04 October 2021

In my view, however, there is a compelling case to be made against endorsing NA and using it to defend FSCs.[26] Here is an example that appears to make trouble for NA:

> **Lightning Sloth**: There are no sharks present to prevent a rescue by John. The evil observer is now monitoring John's thoughts. 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 be immediately struck dead by a bolt of lightning. However, the observer plays no role in causing John's decision to refrain from attempting a rescue.

In this case, it is implausible to hold that the child would not have drowned had John's omission not obtained. Just as it is implausible to claim that John omits to save the child if he tries and fails. So too, it would be very strange to say that John would have omitted to save the child if he had been struck by lightning because he had seriously considered attempting to rescue the child. NA thus yields the result that John is not responsible for the resulting death of the child. But this seems counterintuitive. Lightening Sloth intuitively should receive the same verdict as Sloth. In Sloth, the evil observer would stop John from rescuing the child by getting him to omit to rescue the child. So in Sloth the child would not have drowned had John's omission not obtained. Thus, NA allows for John to be responsible in Sloth. If they endorse NA, the defenders of FSCs will once again have to treat two cases (Sloth and Lightening Sloth) differently when it appears that the two cases should be treated the same.

Consider also the following case offered by Frankfurt (1994):

> **No Movement**: Stanley sits in a room and deliberately refrains from moving his body at all for a period of five minutes. Since Stanley is the only person in the room, his lack of movement entails that no movement occurs in the room. Unbeknownst to Stanley, there is a counterfactual intervener who would have prevented Stanley from moving his body at all had he attempted to do so.[27]

It seems true in this case both that (i) Stanley caused it to be the case that no movement occurred in the room and that (ii) Stanley is responsible for the result that no movement occurred in the room. [Frankfurt (1994) endorsed (ii).] However, it is not true that in the absence of the omission, the result would not have occurred. This is because if Stanley had tried and failed to move he would not have omitted to move but it would still have been true that no movement occurred in the room.[28] Thus, it appears that we have another prima facie counterexample to NA. (Note that I do not wish to claim that Stanley is in

[26] See Clarke (2014) for a different critique of Sartorio's approach.

[27] See Frankfurt (1994: 620–1).

[28] One could reject this claim and hold that one omits to do A if one tries and fails to do A. However, on this understanding of omissions, NA could no longer be used to rule out John's responsibility in Penned-in Sharks. This is because it would then be the case that in Penned-in Sharks John still omits to save the child if he is prevented from doing so by the intervener.

Downloaded from https://academic.oup.com/pq/article/66/264/579/2461022 by William & Mary Libraries user on 04 October 2021

fact responsible. What I am claiming is that those who endorse our intuitive reaction to the FSCs should say that Stanley is responsible.)

In my view, it seems that Stanley is responsible in No Movement, and that John is responsible in Lightning Sloth for the same sorts of reasons that it seems that Jones is responsible in Original Frankfurt Case. So NA apparently does not successfully capture the intuitive difference between cases where the counterfactual intervener appears to be relevant and the cases in which he appears to be irrelevant. As a result, those who accept our basic intuitions about FSCs ought to reject NA, and thus cannot appeal to NA in order to reject (P3).

IV. SARTORIO AND P4

Recall (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.

In my view, (P4) is more intuitively obvious than (P2) and (P3). So I am not inclined to think that a reply to the No Principled Difference argument which relies on rejecting (P4) will be the best way to go. Nevertheless, Sartorio has argued that we have reasons to doubt that John is responsible in Sloth which do not also provide reasons to doubt that John is responsible in Hero. On Sartorio's view, the reason we think that John is responsible in Sloth is that we think that (a) he is responsible for his decision not to jump into the water, and (b) his decision caused the child's death.[29]

Sartorio finds (b) to be questionable. She suggests that 'what caused the child's death is not my *decision not* to jump in, but my *failure to decide* to jump in'.[30] Now, it might seem that this does not matter very much, since we could then say that (a′) John is responsible for his failure to decide to jump in and (b′) his failure to jump in caused the death of the child.[31] But Sartorio also casts doubt on (a′). She points out that you cannot infer from the mere fact that no one forced John to fail to decide to jump in that he is responsible for failing to decide to jump in.[32] Since both (b) and (a′) are questionable, Sartorio concludes that we are left without a good reason to think that John is responsible.

In my view, we do have good reason to accept both (b) and (a′). Consider first (a′). I agree that the mere fact that John was not forced to fail to decide to

[29] Sartorio (2005: 463–4). My (a) and (b) correspond to Sartorio's (1) and (2′)
[30] Sartorio (2005: 464).
[31] My (a′) and (b′) correspond to Sartorio's (1′) and (2″)
[32] Sartorio (2005: 465).

Downloaded from https://academic.oup.com/pq/article/66/264/579/2461022 by William & Mary Libraries user on 04 October 2021

jump in does not entail (a′). But in my view, (a′) is quite intuitive. Note that it is very strange to say that John is responsible for his decision not to jump in but is not responsible for failing to decide to jump in. In normal cases, responsibility for the decision and the corresponding failing appear to stand or fall together (at least when we pick out the act that the agent decided not to do and the corresponding failing under the same description). In addition, I do think that there is a fairly strong argument for (a′). Consider

(c) John is responsible for causing his failure to decide to jump into the water.

And

(d) if John is responsible for causing his failure to decide to jump into the water then he is responsible for his failure to decide to jump into the water.[33]

Now (c) and (d) together entail (a′). Both (c) and (d) appear to be quite plausible. Although in general one need not be the cause of one's failure to decide, it seems plausible to say in this case that *by deciding* not to jump John causes himself to fail to decide to jump in. Furthermore, so long as we are granting that the mere presence of the evil observer does not remove responsibility for decisions, it is hard to see why John would not be responsible for causing himself to fail to decide to jump in.

One might worry that in order to say that John causes his failure to decide, I must posit simultaneous causation since John might both decide not to jump in and fail to decide jump in at the same moment. I'd like to remain neutral regarding the possibility of simultaneous causation, and I see two ways of resolving this short of embracing it. One way would be to claim that while John's decision at t does not cause his failure to decide to jump in at t, it does cause his failure to decide to jump in at t+1. And John's failure to decide at t+1 will be part of the causal explanation of the child's death. The other option would be to couch the argument in terms of explanation rather than causation. Explanation is a broader notion than causation, and there does not appear to be any problem with simultaneous explanation. (For example, certain chemical properties of my coffee cup at t explain why my cup is hot at t.) We would then claim that John's deciding not to jump in provides an explanation of his failure to decide to jump in and that this grounds his responsibility for his failure to decide to jump in.

In addition to (a′), (b) is also plausible. If I am right that by deciding not to jump in the water John causes himself to fail to decide to jump in, then, given that this failure causes the child's death; it would certainly appear that John's decision does play a causal role in bringing about the death of the child. (Again, we could move to the claim that John's decision plays an explanatory

[33] Note that this is just a claim about the Sloth case. It does not purport to be a principle about responsibility in general.

Downloaded from https://academic.oup.com/pq/article/66/264/579/2461022 by William & Mary Libraries user on 04 October 2021

role if that is preferable.) So, in short, (a′) and (b) both appear to be true, and if either of them are true, then Sartorio's argument fails.

As with the other replies we have considered, those who would reject (P4) would need to explain why some mere B-Factors are relevant to moral responsibility while others are not. Why is the counterfactual intervener relevant to John's responsibility in Sloth, but not in Hero? Presumably, Sartorio would attempt to appeal to NA, or some similar principle in order to do this. But as we have seen NA is not very plausible (and also does not appear to yield the needed result in Sloth).

It is worth pausing here to consider Sartorio's (2013) more recent attempt to explain why John is not morally responsible in Sharks. Sartorio defends the following causal principle:

> DM2-CAUSATION: Causes make a difference to their effects in that the effects would not have been caused by the absence of their causes.[34]

Sartorio then reasons as follows:

> By appeal to DM2-CAUSATION, we can similarly explain the lack of responsibility by agents in [cases like Sharks]. In the sharks case, failing to jump into the water doesn't make the relevant kind of difference because, given the presence of the sharks, there isn't the required contrast between what it contributes to the child's death and what its absence (i.e. [John's] act of jumping in) would have contributed to it. [John's] failing to jump into the water, then, doesn't cause the child's death. Hence, it follows from an actual sequence view of responsibility of the kind I am envisaging that [John] cannot be responsible for failing to save the child in virtue of having failed to jump in.[35]

I do not see how it follows from DM2-CAUSATION that *John's failing to jump in* does not cause the child's death in Sharks. The absence of John's failing would clearly not cause the death. So, for all DM2-CAUSATION says, John's failing to jump in may well be a cause of the death. (I'm not claiming that it is a cause, just that DM2-CAUSATION doesn't rule out its being a cause.) So I do not see how DM2-CAUSATION can provide an explanation of John's lack of responsibility in Sharks.

V. SHOULD WE REJECT (P1)?

If the reader has been persuaded by what I have said up to this point, she will be convinced that none of the potential replies to the No Principled Difference argument currently on offer in the literature are successful. Each reply had its own problems. In addition, the *General Problem* still stands. By rejecting either

[34] Sartorio (2013: 193).
[35] Sartorio (2013: 209–210).

Downloaded from https://academic.oup.com/pq/article/66/264/579/2461022 by William & Mary Libraries user on 04 October 2021

(P2), (P3) or (P4) one commits oneself to the claim that mere B-Factors are sometimes relevant to moral responsibility, thus rejecting the principle that appears to underlie our intuitive response to the FSCs. However, there is a potential reply that does avoid the *General Problem*. One could bite the bullet and claim that John is morally responsible for failing to save the child in Sharks. It is to this possible reply that I now turn.

The trouble generated by the No Principled Difference argument is structurally similar to issues that arise in other philosophical contexts. Consider, for example, the trolley problem. We find that there are two cases in which we are inclined to different judgments about the permissibility of an act. We judge that it is permissible to throw a switch in order to redirect a runaway trolley so that it will run over one rather than five. But we also judge that it is not permissible to kill one healthy person in order to harvest his organs and save five lives.[36] Further, reflection on related cases shows us that it is very difficult to find a principled difference which would justify our holding different views about the two cases.[37] Someone convinced that there is no difference with regard to permissibility between the two cases might offer an argument which begins with the trolley case and concludes that it is permissible after all to kill the healthy person in the organ transplant case. However, it is not obvious that this argument would be any better than one which began with the organ transplant case and concluded that it is not permissible to throw the switch in the original trolley case.

This leads us to a potential strategy which a critic of the No Principled Difference argument might employ. This critic could reverse the direction of the argument. Once we see that there is no responsibility-relevant difference between Sharks and Original Frankfurt Case, we ought to conclude, not that Jones is not responsible in Original Frankfurt Case, but rather that John *is* responsible in Sharks. This new argument would still make use of the same (P2), (P3) and (P4) as my original argument. In short, our critic claims, all I have shown is that we must reject either our intuition about Sharks or our intuition about Original Frankfurt Case. I have not shown that it is preferable to reject our intuition about Sharks, and thus, have not shown that critiques of PAP based on appeals to FSCs are unsuccessful.

I agree that this reply is an option for those who wish to avoid the conclusion of the No Principled Difference argument, so let me explain why I think we have more reason to hold onto our intuition about Sharks than we do to hold onto our intuition about Original Frankfurt Case. First, my intuition that John is not responsible in Sharks is initially somewhat stronger than my intuition that Jones is responsible in Original Frankfurt Case. I take it that those who agree with me about the initial weight of the two intuitions will have reason

[36] See Thomson (1976) and (1985).
[37] See, for example, Fischer (1991).

Downloaded from https://academic.oup.com/pq/article/66/264/579/2461022 by William & Mary Libraries user on 04 October 2021

to prefer holding onto their intuition about Sharks. Second, Sharks is a more realistic, down to earth scenario than Original Frankfurt Case. In my view, all other things being equal, we have more reason to trust our intuitions about realistic cases than we do about farfetched cases. I realize that not everyone will agree with this methodological claim, and I do not deny that farfetched cases are evidentially relevant. So I grant that this point may be of limited significance.

Third, my confidence that Jones is responsible in Original Frankfurt Case is diminished when I consider the following question raised by David Widerker (2000): '*what, in your opinion, should* [Jones] *have done instead*?' When I consider Widerker's question I do not lose the intuition that Jones is responsible, but I do find myself feeling less confident that he is responsible. I know of no similar way of reducing my confidence that John is not responsible in Sharks. Thus, I suspect that those who share my reaction to Widerker's question will find it more palatable to reject their intuition about Original Frankfurt Case.

Lastly, if defenders of FSC based critiques of PAP are correct in holding that we should not accept PAP if our intuitions about FSCs are correct, then the intuitiveness of PAP itself provides us with good reason to prefer our intuition about Sharks over our intuition about Original Frankfurt Case. The following principle underlies this claim: if we must reject either A or B, and A is consistent with intuitively plausible principle C but B is not, then, all things being equal, we ought to reject B. Many find PAP to be quite intuitive; it appears that those who do will have reason to hold on to their intuition about Sharks.

In conclusion, I have argued that no one has yet successfully shown that there is a principled distinction to be drawn between Frankfurt-style cases and cases in which the agent is clearly not responsible. Furthermore, there are general reasons for thinking that this task will not be easily accomplished. This result gives us reason to doubt that our intuitions about FSCs are reliable, and thus gives us reason to be sceptical of FSC based critiques of the Principle of Alternative Possibilities.[38]

REFERENCES

Byrd, J. (2007) 'Moral Responsibility and Omissions', *The Philosophical Quarterly*, 57: 56–67.
Clarke, R. (1994) 'Ability and Responsibility for Omissions', *Philosophical Studies*, 73: 195–208.
—— (2011) 'Omissions, Responsibility, and Symmetry', *Philosophy and Phenomenological Research*, 82: 594–624.

[38] For helpful comments on earlier versions of some of this material thanks to Randolph Clarke, D. Justin Coates, Joshua Hollowell, Ben Mitchell-Yellin, Garrett Pendergraft, John Perry, Michael Nelson, Adam R Thompson, Patrick Todd and an anonymous referee. Thanks also to the UCR Agency Writing Workshop and to audiences at the Central APA and the Tennessee Value and Agency Workshop. And extraspecial thanks are due to John Martin Fischer for his significant help and encouragement.

Downloaded from https://academic.oup.com/pq/article/66/264/579/2461022 by William & Mary Libraries user on 04 October 2021

—— (2014). *Omissions: Agency, Metaphysics, and Responsibility*, Oxford: OUP.
Fischer, J. M. (1991) 'Thoughts on the Trolley Problem', in M. Fischer and M. Ravizza (eds) *Ethics: Problems and Principles*, 308—71. Harcourt Brace Jovanovich.
—— (2008) 'Responsibility and the Kinds of Freedom,' *Journal of Ethics*, 12: 203–28.
—— (2010) 'Replies', *Philosophy and Phenomenological Research*, LXXX: 267–78.
—— and R. Mark (1991) 'Responsibility and Inevitability', *Ethics*, 101: 258–78.
—— and —— (1998) *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge: CUP.
Frankfurt, H. (1969) 'Alternate Possibilities and Moral Responsibility', *Journal of Philosophy*, 66: 829–39.
—— (1994) 'An Alleged Asymmetry between Actions and Omissions', *Ethics*, 104: 620–3.
McIntyre, A. (1994) 'Compatibilists Could Have Done Otherwise: Responsibility and Negative Agency', *The Philosophical Review*, 103: 453–88.
Sartorio, C. (2005) 'A New Asymmetry Between Actions and Omissions', *Noûs*, 39: 460–82.
—— (2013) 'Making a Difference in a Deterministic World', *The Philosophical Review*, 122: 189–214.
Swenson, P. (2015) 'A Challenge for Frankfurt-Style Compatibilists', *Philosophical Studies*, 172: 1279–85.
Thomson, J. J. (1976) 'Killing, Letting Die, and the Trolley Problem', *The Monist*, 59: 204–17.
—— (1985) 'The Trolley Problem', *The Yale Law Journal*, 94: 1395–415.
Widerker, D. (2000) 'Frankfurt's Attack on the Principle of Alternative Possibilities: A Further Look', *Philosophical Perspectives*, 14: 181–201.

Rutgers University, USA

Downloaded from https://academic.oup.com/pq/article/66/264/579/2461022 by William & Mary Libraries user on 04 October 2021