



BRILL



brill.com/jmp

Nonresponsible Killers

Jeff McMahan

University of Oxford

jeff.mcmahan@philosophy.ox.ac.uk

1 Introduction

Suppose that a soldier is fighting in a war that is just. His unit is about to be attacked by child soldiers who he knows were earlier forcibly abducted from their homes, brutalized, indoctrinated, and convincingly lied to about their adversaries' aims. Before being sent into combat, these children are also involuntarily administered drugs that suppress their inhibitions about killing people. The soldier believes, let us suppose correctly, that these child soldiers are not only not culpable for the threats they pose, but are not morally responsible for them at all. They are what I will call *nonresponsible threateners*. One of them, who seems to be about 12 years old, has just appeared and will kill the soldier unless the soldier kills him. Is it permissible for the soldier to defend his life by killing the child soldier?

Although both common sense moral intuition and the traditional theory of the just war say that it is permissible for this soldier to kill the child soldier in self-defense, some philosophers have argued that it is not. In 1994, for example, Michael Otsuka and I independently wrote and published essays in which we each argued (1) that it is normally impermissible to kill an innocent bystander in the course of defending the life of another person, (2) that there is no morally significant difference between an innocent bystander and a nonresponsible threatener, and (3) that it is therefore normally impermissible to kill a nonresponsible threatener either in self-defense or in defense of another person ("other-defense").¹ This argument appealed to intuitions about cases but the deeper explanation of the impermissibility of killing in each case is that neither an innocent bystander nor a nonresponsible threatener forfeits the right not to be killed; nor can the killing of either be justified as the lesser evil (unless, perhaps, the bystander or threatener would unavoidably die soon

1 Michael Otsuka, "Killing the Innocent in Self-Defense," *Philosophy and Public Affairs* 23 (1994): 74–94; Jeff McMahan, "Self-Defense and the Problem of the Innocent Attacker," *Ethics* 104 (1994): 252–90.

in any case). But if there is neither a liability nor a lesser-evil justification for the killing, the burden of justification for overriding the constraint against killing an innocent person is not met.

The argument that Otsuka presented is, however, superior to the one I advanced in two respects. First, he restricts his discussion to nonresponsible threateners, whereas I defended the broader claim, which I now reject, that it is normally impermissible to kill an *innocent* threatener, who might be either nonresponsible or morally responsible but not culpable.² Second, Otsuka compared killing a nonresponsible threatener both with killing an innocent person as a *side-effect* of defensive action and with using the killing of an innocent bystander as a *means* of self-preservation, whereas I compared it only with the latter. As we will see, it is crucial that Otsuka's argument includes the comparison with killing as a side-effect.

Helen Frowe believes that it is often permissible to kill a nonresponsible threatener in self- or other-defense. She believes, for example, that it is permissible for the soldier to kill the child soldier who will otherwise kill him. She devotes the first three chapters of *Defensive Killing* to a discussion of nonresponsible threateners, to whom she refers as "innocent threats."³ She begins by seeking to refute Otsuka's argument and then advances a series of considerations that are together intended to show that in certain cases it can be permissible to kill a nonresponsible threatener in defense of only a single person. My aims in this essay are to defend Otsuka's argument, to explain why I find Frowe's positive arguments unconvincing, and to offer further arguments for the claim that it is normally impermissible to kill a nonresponsible threatener in defense of a single victim.

2 Frowe's Critical Argument

Frowe does not dispute the intuition stated in premise (1) of the argument that Otsuka and I gave (that it is impermissible to kill an innocent bystander in the course of defending the life of only one other person), though, as we will see, she interprets it more narrowly than either of us did. She does, however, reject premise (2), which she and Otsuka call the Moral Equivalence Thesis. She does not claim that a nonresponsible threatener has a different moral status from that of an innocent bystander; rather, the moral difference between killing a

2 For elucidation, see Jeff McMahan, *Killing in War* (Oxford: Clarendon Press, 2009), ch. 4.

3 Helen Frowe, *Defensive Killing* (Oxford: Oxford University Press, 2014), pp. 21–87. Her definition of an innocent threat is on p. 21.

nonresponsible threatener and killing an innocent bystander is in the different modes of agency.⁴ Warren Quinn distinguishes between harmful “agency that benefits from the presence of the victim” and harmful agency that “aims to remove an obstacle or difficulty that the victim presents.” And he suggests that “it would not be surprising if we regarded fatal or harmful exploitation as more difficult to justify than fatal or harmful elimination.”⁵ We can, following Frowe, call the first of these forms of agency “exploitative,” and, following Quinn, call the second “eliminative.” According to Quinn’s understanding, both exploitative and eliminative harmings are intended as a means. The difference is that exploitative harming involves *using* the victim, and thus requires her presence, whereas eliminative harming does not.

Frowe embraces Quinn’s plausible suggestion that the constraint against exploitatively harming an innocent person is stronger than that against eliminatively harming an innocent person. And indeed traditional deontological ethics has generally held that the harmful *using* of a person is specially objectionable. I believe, however, that Frowe both overstates and understates her case here.

She understates it by claiming that what is particularly morally objectionable about exploitative killing is that it “manifests the attitude that the innocent person is a tool to be appropriated for gain by others.” “*This*,” she writes, “is what makes [such killings] wrong: ...that [they] exhibit a particularly morally abhorrent attitude towards the person one kills.”⁶ But the constraint against exploitative killing is not primarily a constraint against having a certain attitude. One might have that attitude towards a person but never act on it. That would be objectionable but what is particularly objectionable is *acting* in a way that harmfully and nonconsensually uses a person as a means to one’s ends.

She overstates her case when she argues that

harming a bystander without her consent to avoid anything other than a much greater harm to another innocent person always treats her as a means. This is why it is impermissible to harm bystanders in the course of defending oneself. In contrast, harming a person who poses a threat—innocent or otherwise—does not treat that person as a mere means.⁷

4 Compare McMahan, *Killing in War*, pp. 170–71.

5 Warren S. Quinn, “Actions, Intentions, and Consequences: The Doctrine of Double Effect,” *Philosophy and Public Affairs* 18 (1989): 334–51, p. 344.

6 Frowe, pp. 53–53 (emphasis in the original).

7 *Ibid.*, p. 46. Compare pp. 52, 56, and 58, where there are further discussions of the idea that exploitative killing treats a person as a *mere* means.

Here Frowe seems to acknowledge that what is particularly objectionable about self- or other-preservative action that kills an innocent bystander is not just that it manifests a certain attitude but that it actually harms the person as a means. There are, however, two problems in this passage. One is that the eliminative killing of a nonresponsible threatener also treats that person as a means, as killing her is the defender's intended means of defense. It does not, of course, involve *using* her as a means, but using someone is not the only way of harming her as a means and therefore treating her as a means.

The second problem is that, as Derek Parfit has argued, harmfully using a person as a means does not entail treating that person as a *mere* means, or merely as a means.⁸ In the passage quoted, Frowe implicitly acknowledges that there can be a lesser-evil justification for the exploitative killing of an innocent bystander when that is necessary to save the lives of a sufficiently large number of other innocent people. A well-motivated moral agent who exploitatively kills an innocent person in such conditions treats the victim as a means but not *merely* as a means, as the agent would not have killed the victim as a means of saving fewer innocent people. And the agent might accept substantial cost to himself if that is necessary for the victim to be killed painlessly rather than painfully. That shows that the agent is treating the innocent victim not merely as a means but also as an end, or as someone who matters in her own right.

Frowe's claims about attitudes and treating people as *mere* means are not necessary to her argument and we can ignore them here. Suppose that she, Quinn, and others are right that the exploitative killing of an innocent person is more seriously objectionable, when all else is equal, than the eliminative killing of an innocent person. It is this, she argues, that explains why it is in fact more seriously objectionable to kill an innocent bystander than it is to kill a nonresponsible threatener and thus shows that the Moral Equivalence Thesis is mistaken. And this in turn shows that, even though it is impermissible to kill an innocent bystander in the course of defending one innocent person, it does not follow that it is also impermissible to kill a nonresponsible threatener in defense of one innocent person.

The reason she thinks that the greater wrongness of exploitative killing refutes the Moral Equivalence Thesis is that, on her view, whereas the killing of a nonresponsible threatener is not exploitative, the killing of an innocent bystander *always* is. Otsuka and I assumed, on the contrary, that while the defensive killing of a nonresponsible threatener is indeed eliminative, the killing of an innocent bystander need not be exploitative, but could also be either eliminative or a side-effect of defensive action. We shared the assumption that

8 Derek Parfit, *On What Matters*, Volume One, chapter 9, especially section 31 (pp. 221–28).

the killing of an innocent bystander can be eliminative with Judith Thomson, who had presented an example of a person who, simply by virtue of her location, innocently obstructs a potential victim's ability to escape from a lethal threat from a different source. Those who satisfy this description have come to be referred to as "innocent obstructors" in the literature but Thomson described this person as an innocent bystander.⁹ In later work, I borrowed Thomson's example with the explicit aim of showing that one can kill an innocent bystander eliminatively. I claimed that, if an innocent obstructor, Hiker, were on a narrow, wobbly public bridge that one needed to get across to escape from a culpable threatener, Murderer, one's shaking the bridge to tumble Hiker off, thereby causing her death, would constitute the eliminative killing of an innocent bystander. Intuitively, it would also be an impermissible killing. I then suggested that killing a nonresponsible threatener, which is also eliminative, is no less objectionable, and therefore also impermissible. I also claimed that if one simply ran onto the bridge, merely foreseeing that Hiker would be shaken off (perhaps intending to leap over her if she were to manage to hang on), that would be an instance of killing an innocent bystander as a side-effect.¹⁰ Otsuka's argument explicitly uses an example of killing an innocent person as a side-effect of defensive action, and he refers to the victim in this case as an innocent bystander.¹¹

Frowe argues, however, that none of these is an instance of killing an innocent bystander. She defines a bystander as

a person whose actions, movements, or presence do not endanger Victim [a name or label she uses throughout the book to refer to any potential victim of a threat]. Her presence does not reduce the number of courses of defensive action morally available to Victim. However, harm directed at this person would nonetheless serve to avert or mitigate the threat to Victim.¹²

Frowe plausibly contends that an innocent obstructor contributes to the threat to the potential victim, citing for support the case of a malicious obstructor whose causal role is identical. Obstructors, therefore, cannot be bystanders.

9 Judith Jarvis Thomson, "Self-Defense," *Philosophy and Public Affairs* 20 (1991): 283–310, p. 290.

10 McMahan, *Killing in War*, p. 171.

11 Otsuka, p. 85.

12 Frowe, p. 22.

It is important to notice that Frowe's second criterion of bystander status – that one does not reduce the number of permissible defensive options – is only a necessary condition, not a sufficient condition, of being a bystander. Suppose, for example, that Victim is one of five innocent people fleeing from Murderer who need to get across the bridge to a waiting car that will enable them to drive to safety. All five run onto the bridge intending to jump over Hiker but foreseeing that she will almost certainly be shaken off before they reach her. Suppose that, because this is necessary if all five are to avoid being murdered, the killing of Hiker as a side-effect is proportionate and therefore permissible. It is true of Hiker, then, that “her presence does not reduce the number of courses of defensive action morally available to Victim.” But this does not make her a bystander, for her presence does *endanger* Victim. This is shown by there being a risk that Murderer might be able to kill the other four innocent people by outrunning them. In that case, Hiker's presence would make the defensive option of running across the bridge “morally unavailable” to Victim. Thus, even if Hiker's presence does not in fact reduce Victim's permissible options, it does *contribute* to the threat to his life.

In these cases, in which Victim could leap over Hiker if necessary, Hiker is not a physical obstacle to Victim's ability to avoid being killed. She is not an innocent obstructor but is instead what Gerhard Øverland calls a *moral obstacle*.¹³ According to Frowe, a moral obstacle contributes to a threat, albeit not causally, and thus cannot be a bystander. Indeed, anyone who would be harmed either eliminatively or as a side-effect of defensive action is, in Frowe's taxonomy, not a bystander but a threatener. This clearly conflicts with ordinary usage, as people do say of innocent civilians in war (children, for example) who would be killed as a side-effect of an otherwise permissible attack on a military target, that they are innocent bystanders rather than contributors to whatever threat the attack would be intended to avert. But we can ignore this, as our concerns are moral rather than linguistic.

Frowe goes on to distinguish between what she calls “an indirect threat, ... a person who endangers Victim but who is not going to kill Victim” and a “direct threat,” who, if not prevented, “is going to kill Victim.”¹⁴ Innocent obstructors and moral obstacles are indirect threats – or, as I will say, “indirect threateners.” (Henceforth, unless otherwise indicated, all threateners, direct or indirect, cited in this article should be understood to be in no way responsible for the

13 Gerhard Øverland, “Moral Obstacles: An Alternative to the Doctrine of Double Effect,” *Ethics* 124 (2014): 481–506. For an independently developed but very similar view, see Alec Walen, “Transcending the Means Principle,” *Law and Philosophy* 33 (2014): 427–64.

14 Frowe, p. 22.

threat they pose. Moral obstacles may be necessarily innocent, and perhaps necessarily nonresponsible; for if a person is responsible or culpable for being where he will unavoidably be harmed as a side-effect of otherwise justified defensive action, he is presumptively liable to proportionate harm as a side-effect and thus may not constitute a moral impediment.)

As we will see, Frowe believes that the distinction between direct and indirect threateners is morally significant. But the important point for the moment is that, if a person can be harmed eliminatively, or would be harmed as a side-effect by necessary defensive action, he is, in Frowe's terms, a direct or indirect threatener and therefore cannot be a bystander.

In this taxonomy, Frowe notes, "bystander killings are necessarily *exploitative*. ... But this wrongness can hardly be found in the killing of an innocent threat. Such a killing is eliminative."¹⁵ Yet this may not be quite right. It seems that status as a bystander is relative to a threat, so that a person can be an innocent bystander relative to threat X, which is posed by someone else, and simultaneously be a nonresponsible threatener relative to threat Y, which he poses. Suppose that an agent kills this person intending to defend the potential victim of threat Y but foreseeing that doing so will also have the effect of eliminating threat X. This act of killing would be eliminative but would have an exploitative effect. We can, however, ignore this complication.

Cases of potential rescue raise a further question for Frowe's taxonomy, as it is unclear where potential rescuers fit. Suppose that Victim will die unless Potential Rescuer saves him. Neither is in any way responsible for the threat to Victim. Suppose that Potential Rescuer is permitted but not required to harm herself as a means of saving Victim. She is then an innocent bystander who may permissibly harm herself exploitatively. But suppose that rescuing Victim would involve harming herself only as a side-effect. She then cannot be a bystander in Frowe's taxonomy. But she also does not seem to be a threatener. She is not an obstructor and also cannot be a moral obstacle, as there is, we may suppose, only a prudential rather than a moral reason for her not to harm herself.

Whatever problems Frowe's conceptual and taxonomical proposals might raise, the argument she builds on the basis of her definition of a bystander is that (1) all killings of innocent bystanders in the course of defensive action are necessarily exploitative, whereas (2) killings of nonresponsible threateners are eliminative (or side-effects, though she does not say this), and (3) because killing an innocent person exploitatively is more seriously objectionable than killing an innocent person eliminatively, (4) the Moral Equivalence Thesis is

¹⁵ Ibid., p. 8.

therefore false and “Otsuka’s argument that the killing of an innocent threat can be morally assimilated to the killing of a bystander is mistaken.”¹⁶ This leaves it open that it can be permissible to kill a nonresponsible threatener in defense of a single person even though it is impermissible to kill an innocent bystander to preserve the life of a single person.

This argument, I believe, misses its target, as it is concerned only with how we might or should use words. The substance of Otsuka’s argument can be restated in Frowe’s terms without losing any of its force in the translation. Here is a suggested reformulation of the basic argument. It differs from Otsuka’s original argument in certain ways that I think strengthen that argument. Most notably, unlike the original argument, it makes no reference to innocent bystanders in Frowe’s sense. Otsuka not only stated his initial premise with reference to bystanders but also illustrated it with an instance of the exploitative killing of a bystander. This, I believe, was a mistake. Innocent bystanders in Frowe’s sense are irrelevant to the argument that he and I were attempting to develop.

The first premise in that argument is that the eliminative killing of an indirect threatener, such as a nonresponsible obstructor, is normally impermissible as a means of defending only a single person, as is the killing of a moral obstacle as a side-effect of defending the life of only one person. (One might include a proviso that the harm of death to the indirect threatener would not be slight, or very substantially less than the harm of death to the person who might be defended.) That Otsuka and I earlier referred to such indirect threateners as innocent bystanders makes no difference to the intuition that it is impermissible to kill them. Indeed, Frowe herself accepts the claim of this premise, just as she accepts the corresponding premise of Otsuka’s original argument, provided that the concept of an innocent bystander as it appears there is understood in her sense. Yet, understood in her way, the first premise of the original argument, which refers to innocent bystanders, is quite different in substance from the first premise of this revised and restated argument.

The second premise is that there is no difference between the eliminative killing of a direct threatener and either the eliminative or collateral killing of an indirect threatener that makes killing a direct threatener less morally objectionable or easier to justify than killing an indirect threatener. The conclusion of the argument is then that the killing of a direct threatener in defense of a single victim is normally impermissible.

16 Ibid.

This argument can be illustrated by examples drawn from an excellent recent article by Amir Saemi and Philip Atkins, though relabeled here and presented with somewhat different details.¹⁷

Shield Behind Murderer is about to kill Victim. Victim – an American and thus equipped with a high-powered gun – can kill Murderer in self-defense but his bullet will pass through Murderer’s body, killing a nonresponsible person, Shield, who is behind him.

Shield in Front As in Shield Behind except that Shield is in front of Murderer so that Victim’s bullet must pass through her body before killing Murderer.

Obstructor As in Shield in Front except that Victim has been insufficiently attentive to the urgings of the National Rifle Association and thus has only a low-powered gun. To defend himself against Murderer, he must first shoot Shield to get her out of the way and then shoot Murderer. The only parts of Shield’s and Murderer’s bodies that are exposed are their heads, so shooting Shield will unavoidably kill her.

Suppose Victim kills Murderer as a means of self-defense in Shield Behind. He also unintentionally kills Shield as a side-effect. This side-effect killing is causally “downstream” from Victim’s means of saving his life. Shield is here a moral obstacle, as her presence behind Murderer constitutes a moral reason for Victim not to shoot. Thomson, Otsuka, and I have referred to people like Shield as innocent bystanders but nothing in the argument depends on the use of that term.

Suppose next that Victim kills Murderer in Shield in Front. Here too he kills Shield as a side-effect. He acts exactly as he would if Shield were not there. He does not *use* Shield as a means of getting his bullet to hit Murderer.¹⁸ Nor does he kill Shield as a means of eliminating an obstacle that Shield presents to the killing of Murderer. It is simply that, if it is to kill Murderer, his bullet must pass through Shield’s body, just as it must in Shield Behind. The difference is

17 Amir Saemi and Philip Atkins, “Targeting Human Shields,” *The Philosophical Quarterly* 68 (2018): 328–48, p. 330.

18 Jonathan Quong would argue that Victim uses the space that Shield occupies and to which Shield thus has a right and that this makes the killing exploitative. See his “Killing in Self-Defense,” *Ethics* 119 (2009): 507–37. Frowe criticizes this view on pp. 56–73, as does Michael Otsuka in “The Moral Responsibility Account of Liability to Defensive Killing,” in Christian Coons and Michael Weber, eds., *The Ethics of Self-Defense* (New York: Oxford University Press, 2016): 51–68.

that the killing in Shield in Front is a causally “upstream” side-effect of killing Murderer, whereas the killing in Shield Behind is downstream. In this case, as in Shield Behind, Shield is a moral obstacle. She is not a nonresponsible obstructor.

Shield is, however, a nonresponsible obstructor in Obstructor. When Victim kills Shield in Obstructor, he intends to affect her in a way he knows will kill her and he does this as his means of enabling himself to kill Murderer. Although he kills her as a means, he does so eliminatively rather than exploitatively.

There might, of course, be circumstances in which it is necessary for Victim to kill Shield exploitatively, as in:

Literal Shield Victim has paid no attention to the NRA and has only a knife. The only way he can avoid being shot by Murderer is to subdue Shield by killing her with his knife and then using her body as a shield against Murderer’s bullets.

If Victim kills Shield in this case, he harmfully *uses* her as a means. But such cases are irrelevant to the argument that Otsuka and I have sought to defend, which aims to show that there are instances of defensive killing that are intuitively or arguably impermissible but *no more* morally objectionable than the killing of a nonresponsible threatener in self- or other-defense, so that the killing of a nonresponsible threatener should be impermissible as well. Since killing a nonresponsible threatener in self- or other-defense is eliminative, the exploitative killing in Literal Shield is *more* morally objectionable and thus irrelevant for our purposes.

The first three of these four cases involving the killing of Shield are, I believe, of the relevant sort. While it may not be intuitively obvious that it is impermissible for *Victim* to kill Shield in Shield Behind, it does seem obvious that it is impermissible for an *impartial third party* to kill Shield as a downstream side-effect of killing Murderer. That is because in this case there is no agent-neutral justification, such as a liability justification or a lesser-evil justification, that can override the constraint against killing one innocent person to save another. This leaves it open that Victim might, as Quong argues, have an “agent-relative permission” to give some priority to his own life by killing Shield. Unless there are such permissions, the prohibition that applies to third parties must apply to Victim as well. I cannot pursue this here. For present purposes, it is sufficient that Frowe argues, and I agree with her, that the self-defensive killing of a nonresponsible threatener is not justified by an agent-relative permission.¹⁹

19 Frowe, pp. 54–63.

It seems clearer intuitively that it is impermissible for Victim to shoot through Shield in Shield in Front. The killing of a nonresponsible person as a causally upstream side-effect of an effort to achieve an end seems harder to justify than the killing of a nonresponsible person as a downstream side-effect. I will not attempt to defend this intuition here but will simply assume, for the purpose of argument, that it is defensible.

Finally, it is even more intuitively compelling that it is impermissible to kill Shield in Obstructor than it is that it is impermissible to kill her in Shield in Front. This is explained and, I think, justified by the common view that the killing of a nonresponsible person as an intended means, even eliminatively, is harder to justify than killing a nonresponsible person as a side-effect, even when it is causally upstream rather than downstream. But the killing of a direct threatener in self- or other-defense is also the eliminative killing of a nonresponsible person as an intended means. Unless it differs in some important respect from the eliminative killing of a nonresponsible obstructor, it must be impermissible as well. Frowe argues that there is such a difference – namely that a direct threatener will otherwise *kill* Victim, whereas a nonresponsible obstructor will not.²⁰ I will consider this response at length in section 4 below.

In a recent essay, Otsuka suggests that one should not, in trying to defend the Moral Equivalence Thesis, appeal to cases in which killing is upstream, citing an objection I made earlier to his use of a case of upstream killing in an argument in his 1994 article.²¹ In the recent essay, he instead argues that killing a nonresponsible threatener in self-defense is relevantly like killing a nonresponsible person who occupies a tiny alcove into which one must get to avoid being killed by an approaching train (a case known in the literature as *Alcove*). One might either squeeze into the alcove, crushing the occupant, or take her place by ejecting her, causing her to be killed by the train. Either would be the

20 This same claim is differently defended in Tyler Doggett, “Killing Innocent People,” *Noûs* 52 (2018): 645–66.

21 Otsuka, “The Moral Responsibility Account of Liability to Defensive Killing,” pp. 55–57. On p. 54 of this essay, Otsuka politely but correctly notes a serious error I made in the same discussion of his 1994 article. I think of these disputes between Otsuka and me as our continuing joint efforts to formulate the basic argument that we both still accept in its best form.

I am much indebted to an anonymous reviewer for this journal who called my attention to Otsuka’s essay and the challenge it poses to my appeal to cases of killing as a side-effect. Although I have resisted this reviewer’s suggestion to omit these cases, since I think that Otsuka and I are ultimately in agreement, the reviewer’s careful and constructive comments nevertheless prompted me to rethink my arguments and to make substantial revisions.

killing of an innocent obstructor but not, Otsuka says, a killing that is upstream from “the good of the elimination of the threat to your life,” for the killing and the elimination of the threat are, he says, the same event.²² But my criticism of his earlier argument was that he sought to support the claim that it is impermissible to kill a nonresponsible person as a downstream side-effect (as in *Shield Behind*) by appealing to the analogy with a more clearly impermissible killing of a nonresponsible person as an upstream side-effect. Otsuka is right to concede now that “causally upstream killings appear, intuitively, to be morally more problematic than otherwise similar killings that are not causally upstream.”²³ But this is compatible with using a case of killing as an upstream side-effect to support the Moral Equivalence Thesis. This is because killing a nonresponsible obstructor or moral obstacle as an upstream side-effect is, precisely in being upstream, relevantly analogous to killing a direct threatener as a means of eliminating the threat she poses. Even though the former is killing as a side-effect and the latter is killing as a means, both are killings that are causally upstream from the intended end of self-preservation. Indeed, unless there are instances of noncausal means, all killings that are means, whether exploitative or eliminative, are causally upstream from their ends. This is true of the killing of the occupant in *Alcove* as well. The end in that case to which killing the occupant is a causal means is not “the elimination of a threat to your life” but your continuing to live. “Elimination of a threat” is just a re-description of the means to this end. Otsuka goes on to acknowledge something close to this and rightly claims that his appeal to *Alcove* remains effective even if the killing of the occupant is a causally upstream effect. Both killing the occupant and shooting *Shield* in *Obstructor* are instances of the eliminative killing of a nonresponsible obstructor. There is no morally significant difference between these two killings.

The conclusion of Otsuka’s argument appealing to *Alcove*, and of my appeal to *Obstructor* in the reformulation of his earlier argument, is that the eliminative killing of a direct threatener is impermissible because it is *no less* seriously wrong than the eliminative killing of a nonresponsible obstructor, which is intuitively impermissible. But the reformulated argument supports a stronger conclusion. It seems obvious that it would be impermissible for a disinterested third party to defend *Victim* by killing *Shield* in *Shield in Front*. If, as Frowe and I believe, *Victim* is not permitted to kill *Shield* simply on the basis of partiality toward himself, it must be impermissible for *Victim* to kill *Shield* in *Shield in Front*. But killing *Shield* in *Shield in Front* is an instance of killing as

²² Ibid., p. 56.

²³ Ibid.

a side-effect, albeit an upstream side-effect, whereas killing a direct threatener is killing as an intended means, which is a more objectionable mode of agency by which to kill a nonresponsible person. It follows, if other things are equal, that killing a nonresponsible threatener is *more seriously objectionable* than killing Shield in Shield in Front, which is itself impermissible. It seems to me, moreover, that the fact that the side-effect killing in Shield in Front is causally upstream would not on its own be sufficient to make that killing impermissible if the downstream killing in Shield Behind were permissible; therefore the latter should be impermissible as well, as I claimed earlier.²⁴ But if the killing in Shield Behind is impermissible, the killing of a nonresponsible threatener should be *substantially more objectionable*, and harder to justify, assuming that killing as an intended means is substantially harder to justify, if other things are equal, than killing as a downstream side-effect.

Otsuka makes a similar point in his 1994 article but I think he overreaches. He notes that the defensive killing of a direct threatener is killing as an intended means. But, failing (as I also did at the time) to distinguish between the exploitative and eliminative forms of killing as a means, he assimilates the latter to the former in claiming that killing a direct threatener is “analogous to those most deplorable cases in which you kill a Bystander in order to eat her body to prevent yourself from starving or in order to replace your failing vital organs with her healthy ones.”²⁵ These are instances of the exploitative killing of an innocent person, which, as we have seen, Quinn, Frowe, and other non-consequentialists believe to be more objectionable, and harder to justify, than the eliminative killing of an innocent person.

One might object to the claim that eliminative killing is more objectionable than killing as a side-effect by observing that the common sense intuition seems to be that killing an innocent person as a side-effect of defensive action is normally impermissible while killing a nonresponsible threatener is normally permissible. I believe, however, that the intuition, which I share, that killing a nonresponsible threatener is permissible is a result of at least two mistakes. One is a subconscious conflation of permission with excuse. The other is a tendency to overgeneralize intuitions about the permissibility of defensive killing that we have about far more common cases involving culpable threateners.

Frowe explicitly rejects the second premise of the restated argument presented on page 658, claiming that a direct threatener has an enforceable duty to bear costs to avoid killing an innocent person that is stronger than any

24 For a cogent defense of the claim that there is no morally significant difference between the killing in Shield Behind and that in Shield in Front, see Saemi and Atkins.

25 Otsuka, p. 87.

corresponding duty that an indirect threatener might have to avoid nonresponsibly contributing to a killing. I will examine her argument in Section 4 and offer a defense of the second premise in Section 5. But first I will consider another argument she gives that, like Otsuka's earlier argument, is based on a sequence of examples, each of which seems to differ only trivially from the one that precedes it.

3 Frowe's Intuitive Argument

I noted in Section 1 that in addition to her critique of Otsuka's argument, Frowe presents a series of considerations that she believes support the permissibility of killing direct threateners. These can be separated into two categories, intuitive and theoretical.

Frowe's intuitive argument begins with a hypothetical example that first appeared in Nozick's *Anarchy, State, and Utopia* and has subsequently been discussed by Judith Jarvis Thomson and many others. Frowe calls it

Ray Gun Falling Person is blown by wind down a well at the bottom of which Victim is trapped. Falling Person will crush Victim to death unless Victim vaporizes her with his ray gun. If he does not vaporize her, Victim will cushion Falling Person's landing, saving her life.²⁶

This is intended to be a paradigm case of a direct threatener. It is intuitive that Victim may permissibly kill Falling Person in self-defense. Otsuka and I sought to challenge that intuition by presenting cases that we claimed were no different morally but in which killing is impermissible. Frowe, by contrast, seeks to reinforce the intuition by presenting cases that she claims are no different morally but in which it is even more obvious that killing is permissible. Just as she argues that our cases are not relevantly similar, so I will argue that hers are not either.

Frowe cites four further examples of the same type of situation.

Spacious Well If Victim remains stationary, he will be crushed and Falling Person will survive. If he moves aside, he will survive but she will be killed.

26 This statement of the example is from Frowe, p. 22. Also see Robert Nozick, *Anarchy, State, and Utopia* (New York: Basic Books, 1974), pp. 34–35. The details of Thomson's version are different – for example, Falling Person is caused by a villain to fall toward Victim – but the relations among the persons involved are meant to be the same. See “Self-Defense,” pp. 287ff.

Stationary Shield If Victim remains on top of a shield, he will be killed and Falling Person will survive. If he gets underneath the shield, he will survive but she will be killed.

Existing Shield Victim is underneath a shield. If he removes it, he will be killed and Falling Person will survive. If he leaves it in place, he will survive but she will be killed.

Flagpole Victim is holding a flagpole upright. If he lowers it, he will be killed and Falling Person will survive. If he keeps it in place, he will survive but she will be killed.²⁷

Frowe cites with approval an analysis of the concept of killing given by Otsuka according to which killing consists in “the initiation or sustaining of, or the insertion of somebody into, a sequence of events that results in the death of a person.”²⁸ Following Otsuka, she claims that Victim’s action in *Flagpole* “is a killing because, in continuing to hold the flagpole, Victim sustains a sequence of events that results in the death of Falling Person. ... But of course, *Flagpole* is just *Existing Shield* with a different object. If *Flagpole* is a killing, so too is *Existing Shield*.” Finally, she claims that Victim’s moving aside in *Spacious Well* is also an instance of killing, for it “initiates a sequence of events of which Falling Person’s death is an upshot, since he initiates moving out of her path when so moving will result in her death.”²⁹ That Victim’s staying underneath the shield in *Existing Shield* is morally indistinguishable from his moving aside in *Spacious Well* is reinforced by the intermediate case, *Stationary Shield*, which is like *Existing Shield* except that Victim has to move, as in *Spacious Well*, to avoid being crushed.

Frowe’s point seems to be that, because it is *obviously* permissible for Victim to move aside in *Spacious Well*, and because no morally significant differences are introduced in the sequence of cases between *Spacious Well* and *Ray Gun*, it must be permissible for Victim to kill Falling Person in the latter case as well.

Frowe considers the objection that “this just shows that Otsuka’s account of what it is for something to be a killing needs revising,” but she rejects this suggestion on two grounds. The first is that “Otsuka is correct that intentionally sustaining a lethal sequence of events falls on the killing side of the killing/letting die distinction. Even if Victim does not take up the flagpole with lethal

27 Frowe, pp. 64–67. These are my own statements of the examples. *Flagpole* is taken from Otsuka, p. 89.

28 Frowe, p. 65.

29 *Ibid.*, pp. 65–66.

intent, his continuing to hold the flagpole renders his action more than a letting die."³⁰

Frowe appears to assume that if Victim intends that Falling Person should die, he must intend to kill her. But one's intention is, except perhaps in certain unusual cases, irrelevant to whether one's act is a killing or an instance of letting a person die; for one can intend to allow a person to die, which is in fact what Victim does in *Flagpole*. To show this, it may help to borrow a strategy from Frowe. To make her case that obstructors and moral obstacles are threateners, she seeks to elicit our intuitions in cases in which these people intentionally and culpably act as obstructors or obstacles.³¹ We can similarly suppose that Falling Person has intentionally and maliciously thrown herself from a height intending to crush Victim without injuring herself. But she failed to perceive that he was holding up a flagpole. In this case, if Victim knows what Falling Person has done and fails to lower the flagpole, he simply fails to save her by making himself into a cushion for her, and *she* then impales *herself* on his flagpole. This is even clearer if Victim does not see her descending toward him. In this case he does not kill her, but neither does he allow her to die, as one can allow a person to die only if one is aware that one can save her. One can, of course, kill a person without knowing one is doing so. And this is indeed a case of accidental killing, not because Victim kills Falling Person but because Falling Person accidentally kills herself.

Similar remarks apply to Frowe's three other cases. As she rightly notes, *Existing Shield* is exactly like *Flagpole* except that Victim is protected by a different object. So that too is a case of letting Falling Person die. That Victim's action is not a killing is most obvious in *Spacious Well*. Imagine a variant of that case in which Victim is initially away from where Falling Person will land. If he does not move, he will clearly allow her to die. But suppose that he decides to save her and moves underneath her, then changes his mind and moves back to where he was, so that she hits the ground and dies. He does not kill her but initiates action to save her and then, before the action becomes effective, withdraws what would have saved her and thus allows her to die.³² Yet once Victim has positioned himself where Falling Person will land, this case is from then on exactly like *Spacious Well*. In that case too Victim refuses to save Falling Person by providing his body as a cushion for her. Again, this is even more obvious if we suppose that Falling Person has maliciously hurled herself at him. In that

30 Ibid., p. 66.

31 Ibid., pp. 25–26.

32 For elucidation and defense of this claim, see Jeff McMahan, "Killing, Letting Die, and Withdrawing Aid," *Ethics* 103 (1993): 250–79.

case, although she attempts to kill him, she accidentally kills herself instead by throwing herself down a well in the mistaken belief that he, or his body, would save her.

It is revealing that in a couple of passages Frowe herself describes Victim's action in cases of this sort as a failure to save. She writes that "those who oppose a permission to kill innocent threats must resist a requirement for Victim to rescue innocent threats [such as Falling Person] at the cost of his own life."³³ This presupposes that in the four cases that are supposed to be morally just like *Ray Gun*, Victim *fails to rescue* Falling Person. And this implies that he does not kill her, as it would be absurd to describe an act of killing as a failure to rescue one's victim from being killed by oneself. (There are, admittedly, cases in which one kills a person as a means of preventing her from saving herself at the expense of one's own life. If you and I both need a respirator to survive and I shoot you to prevent you from stealing mine, I both kill you and prevent you from saving yourself. I think that *Ray Gun* is also a case of this sort. But these are not best understood as cases in which one both kills a person and allows that person to die; for preventing a person from being saved is arguably different both conceptually and morally from allowing a person to die.)³⁴

Frowe's second response to the objection that the analysis of killing that she takes from Otsuka's article is mistaken is that

It doesn't matter whether, for example, holding the flagpole is a killing but moving the shield over oneself is a letting die. What matters is whether it is plausible to grant Victim a permission to move the shield over himself, or move himself under the shield, but deny him a permission to continue to hold the flagpole. And I don't see how it can be plausible, because picking up a lethal object – which is what the shield is – looks *more* like a killing than does merely continuing to hold the flagpole.³⁵

This response, however, underestimates the force of the objection, which is that, whereas disintegrating Falling Person in *Ray Gun* is an instance of killing a direct threatener, *all* of the other four cases are instances of allowing a direct threatener to die. And one might reasonably think that Victim's action in these four cases is permissible precisely because they are instances of allowing an

33 Frowe, p. 71. She makes the same claim on p. 66.

34 See Matthew Hanser, "Killing, Letting Die and Preventing People from Being Saved," *Utilitas* 11 (1999): 277–95.

35 Frowe, p. 66.

innocent, nonresponsible person to die rather than instances of killing an innocent, nonresponsible person.

As I indicated earlier, I accept that in these four cases, Victim allows Falling Person to die *intentionally*, as a means of saving his own life. Yet, like intentionally killing a person, intentionally letting a person die can be either eliminative or exploitative. In the four cases, Victim's allowing Falling Person to die is eliminative and thus does not involve harmfully using her as a means. If Victim were to allow an innocent person to die exploitatively – for example, in order that her organ could be transplanted into his body – that might well be impermissible (though it might be permissible if, as in the earlier variants, Falling Person had culpably hurled herself at Victim).

4 Frowe's Theoretical Argument

After advancing her intuitive argument that appeals to the four cases, Frowe develops a more theoretical argument with a number of elements to it. Here I will attempt to summarize that argument and to distinguish its six different elements. Immediately after stating her points 3 and 5, I will briefly register my skepticism about each but will not pursue the matter in depth. After stating all the elements of the argument, I will advance some larger objections.

(1) Unless prevented from doing so, a direct threatener will directly harm or kill Victim. Frowe often adds that the direct threatener will do this by interfering with Victim's body. As it is not clear to me why that feature of the killing is significant, I will ignore it.

(2) Because a direct threatener will otherwise kill Victim, defensive agency directed against her is eliminative and therefore does not involve harmfully *using* her. But this is not, as Frowe acknowledges, sufficient for the act of killing to be permissible. (The killing of a nonresponsible obstructor is not exploitative either, but Frowe claims that it is impermissible to kill an indirect threatener unless she is morally responsible for her contribution to a threat.) That the mode of agency in killing a direct threatener is not exploitative shows only that one possible objection to an act of killing does not apply in this case.

(3) We have a duty not to kill and also a secondary duty to "bear greater costs to avoid killing (or directly harming) people than to avoid endangering them in other [that is, indirect] ways."³⁶ This is the basis of Frowe's claim that there is a morally significant difference between a direct threatener and an indirect

³⁶ Frowe, pp. 78–79.

threatener, and hence of her rejection of the second premise in my revised restatement of Otsuka's original argument.

Frowe here appeals to the view, introduced by Frances Kamm and later developed by Victor Tadros, that if, for example, Falling Person could alter her trajectory so that she would not crush Victim, she would have a duty to do that, at a proportionate cost.³⁷ Kamm thinks, however, that she is not required to bear a lethal cost to do so.³⁸ Frowe agrees in cases in which a direct threatener, such as Falling Person, does not *intend* the threat she poses. But she claims that when a nonresponsible threatener "has intentionally caused herself to threaten, ... her duty to prevent herself from killing Victim requires her to bear lethal costs."³⁹

I simply note here that I see no reason why a threatener's intention should have so much significance, or indeed any significance at all, if she is in no way responsible for having it. A person's intentions have no more moral significance than the movements of her body have if she is not in any degree responsible for either.

(4) The duty to bear costs to avoid killing an innocent person is enforceable by the threatened person and by third parties. Assuming that Falling Person can do nothing to avoid killing Victim, Victim or a third party is permitted to engage in defensive action that imposes costs on Falling Person up to those she would be required to impose on herself to prevent herself from landing on Victim.

(5) Part of the explanation of why we have this enforceable duty not to kill, *even nonresponsibly*, is that we are responsible for whatever our bodies cause to occur. Here Frowe quotes Victor Tadros, who writes that "it is the fact that I am responsible for what my body does, even when it is not a product of my agency, that gives rise to the permission to harm innocent attackers and innocent threats. And I bear that responsibility because my body is *me*."⁴⁰

It is unclear, however, how one could bear no responsibility for the threat one's body poses, as Frowe says of Falling Person (who is, by hypothesis, a *nonresponsible* threatener), and yet simultaneously be "responsible for what [one's] body does," as she follows Tadros in claiming is also true of Falling

37 Frowe cites F.M. Kamm, *Creation and Abortion* (Oxford: Oxford University Press, 1993), p. 47, but Kamm's argument also appears earlier in her "The Insanity Defense, Innocent Threats, and Limited Alternatives," *Criminal Justice Ethics* 6 (1987): 61–76, pp. 64–65. Also see Victor Tadros, *The Ends of Harm* (Oxford: Oxford University Press, 2011), pp. 254–55.

38 Kamm, *Creation and Abortion*, p. 48.

39 Frowe, p. 83.

40 Frowe, p. 68; Tadros, p. 255.

Person. Perhaps what she means is that, although a nonresponsible threatener has no control over the threat her body poses, she is nonetheless picked out as the one to whom the harms it causes must ultimately be distributed, in much the same way that a military officer might be required to bear costs, either as punishment or as compensation to victims, for harms wrongly inflicted by one of her soldiers over whose wrongful acts she could not have exercised any control. In most such cases, however, the officer is held responsible on the ground that there really *was* something she failed to do in the past that might have motivated the soldier not to act as he did, or perhaps something she did that might have motivated him to act in that way. But in this respect there is no parallel with the relation between Falling Person and her body, as we are assuming that there was nothing she could have done to prevent her body from becoming a threat. In other cases, an officer may be held accountable for purely instrumental reasons, such as to provide other officers with an incentive to exert as much control as possible over their soldiers. But again there is no parallel with Falling Person and other similar direct threateners, for part of what it is to be a nonresponsible threatener is that there is nothing one could reasonably have been expected to do to avoid becoming one. If one could have avoided becoming a threatener by taking reasonable precautions, one is to some extent responsible for having become one.

(6) In cases in which a direct threatener has an enforceable duty to bear only less than lethal harm, it can still be permissible for Victim or a third party to kill her. This is because the greater part of the harm in the direct threatener's being killed can be justified on the ground that she has a duty to bear it, while the remaining part can be justified because her suffering it is the lesser evil – that is, substantially less bad than Victim's being killed by her.

In presenting this multi-faceted argument, Frowe claims that a direct threatener's having an enforceable duty to bear defensive harm is not equivalent to her being liable to be harmed. "She is," Frowe contends, "not *liable* to defensive harm—she hasn't *forfeited* her rights. But she nonetheless lacks a right not to be harmed because in harming her...Victim imposes only costs that she has a duty to bear."⁴¹ This seems a distinction without a difference. For simplicity of exposition, assume that the harm that the direct lethal threatener has a duty to bear is the harm of being killed. Frowe affirms that the threatener lacks a right not to be killed. Yet the threatener had this right before she began to pose a lethal threat. The claim that by posing the threat she acquires an enforceable duty to bear a lethal harm entails that by posing that threat she loses her right not to be killed; for it makes no sense to suppose that she has a right against

⁴¹ Frowe, pp. 83–84.

the justified enforcement of her enforceable duty. But losing that right seems no different in substance from forfeiting it, which is what it is to become liable to be killed in conditions in which some harm is unavoidable (unless the forfeiture is a corollary of deserving to be killed, which I will assume it is not).

Yet duty and liability are not strictly correlative. While all those who are liable to suffer a certain harm have a duty to bear it, not all those who have a duty to bear a harm are liable to suffer it. For example, a person who happens by chance to be where an accident occurs can acquire a duty either to harm herself or to allow herself to be harmed by a third party as a means or side-effect of the rescuing of the victim without being morally liable to be harmed in that way. We might say that this person, whom we met earlier as Potential Rescuer, *loses* her right not to have costs imposed on her but does not *forfeit* it, whereas a threatener who is liable to defensive harm does forfeit her right not to suffer that harm.

The question is whether the reason why a person acquires a duty and loses a right by becoming a direct threatener is relevantly similar to or different from the reason why the person at the accident acquires a duty and loses a right. It seems that Frowe should think that these reasons are different. This is because she thinks that the strengths of the two duties are different.

The duty to bear costs in the rescue of an accident victim is grounded in those costs' being the lesser evil – that is, the harm the person has a duty to incur in the saving of the victim is substantially less severe than that which would thereby be prevented. Yet Frowe believes that the harm that a direct threatener has a duty to bear to avoid killing an innocent victim is substantially greater than that which the person at the accident is required to bear in the saving of a life. Otherwise there would be no difference in strength between the duty to prevent one's own body from killing an innocent person and the duty to prevent someone else's body, or some other object, from doing so. This is why a direct threatener's supposed duty to bear costs necessary to prevent her from killing an innocent person should, for Frowe, be correlative with liability rather than being like a person's duty to suffer harm in the saving of an accident victim.

Yet Frowe believes, rightly in my view, that responsibility is necessary for liability. She thus contends that one's becoming a nonresponsible threatener does not make one liable to defensive harm. I believe, similarly, that one's becoming a nonresponsible threatener does not generate an enforceable duty to bear harm that is as great as that which one would be required to bear if one were liable to defensive harm. Nor does it generate an enforceable duty to bear harm that is as great as Frowe thinks it is. Someone who has become a direct threatener without being responsible for doing so does of course have a duty, if

possible, to bear some costs to prevent herself from killing an innocent victim. But the harm she has a duty to bear is, I believe, no greater than that which an innocent, uninvolved person would have a duty to bear if by bearing it she could otherwise harmlessly prevent a direct threatener from killing an innocent victim. In short, if a direct threatener has a duty to bear harm to prevent herself from otherwise *nonresponsibly* killing an innocent person, it is because her bearing that harm is the lesser evil in the circumstances. There is, I believe, no basis for a duty to bear harm that lies between a duty to suffer the lesser evil and a duty that is correlative with liability.

It will help at this point to distinguish among the different types of duty a threatener or potential threatener might have. First, the basic duty is not to engage in voluntary action for which one is morally responsible that will kill an innocent person without justification. But our concern is with duties that apply to or are possessed by those who threaten innocent people nonresponsibly. There is then a second type of duty to take reasonable precautions against becoming a nonresponsible threatener – for example, in the case of Falling Person, a duty, if possible, to avoid being in a situation in which one might be blown or thrown down a well in which a person is trapped. And like many other duties, this duty may entail a secondary duty to bear costs that are necessary for fulfilling it. But again, both these duties are those of morally responsible agents.

Third, as Kamm and Tadros have noted, there seems to be a duty to try to avoid killing an innocent person, and to bear necessary costs in doing so, even after one has become a direct threatener – for example, after one's body has been hurled at another. Again, if Falling Person could alter her trajectory so that she could avoid killing Victim without incurring an unreasonably high cost, she would have a duty to do that.

Yet if someone with the capacity for morally responsible agency poses a direct threat to the life of another innocent person and has the option of preventing herself from killing that person, she is not a nonresponsible threatener, even if she is not responsible for having become a threat. Suppose, for example, that Falling Person has been hurled at Victim but can avoid killing him and ought to bear the costs necessary to do so. Yet again, this is a duty that applies to her as a morally responsible agent. Suppose she refuses to bear those costs and continues to threaten him. In that case, her wrongful refusal makes her *liable* to defensive action that may inflict on her not only the harm she was morally required to bear but also additional harms that may have become necessary and that are also proportionate, taking into account that the amount of harm it is proportionate to inflict on her has increased as a result

of her wrongdoing. She becomes liable because she is responsible for allowing herself to threaten Victim without justification.

None of the duties I have cited is a duty of a nonresponsible threatener. Suppose that Falling Person lacks any ability, either before or after she is hurled down the well, to avoid killing Victim and that she does indeed land on him and kill him. She has not violated any duty. In killing Victim, she has not done or *willed* or even *intended* anything. There cannot, moreover, be a duty not to kill an innocent person wholly involuntarily and nonresponsibly. There cannot, for example, be a duty not to be hurled involuntarily against another innocent person (though, as noted, there can be a duty to resist being hurled in this way). This is an implication of the claim that “ought” implies “can.”

This raises the question of what duty would be enforced if Victim or a third party were to disintegrate Falling Person as she descends. I have claimed that it cannot be a duty that she actually has and will otherwise violate. Perhaps it is a duty that she *would* have *if* she were not a nonresponsible threatener, such as the duty not to do what would kill Victim or the duty to prevent herself from killing Victim, even at some personal cost. But because those are not duties that she violates by being hurled involuntarily at Victim, it seems that they are not duties that can be enforced by killing her. That is, she cannot be made to comply with such duties by being killed if they are not duties that, in the circumstances, she actually has.

Let us assume, however, that I am wrong about this and that if Victim or a third party kills Falling Person, he or they would be enforcing Falling Person's duty not to kill Victim. As my brief summary of Frowe's argument indicates, she follows Kamm in accepting that Falling Person does not have a duty to bear *lethal* defensive harm. The question thus arises of how much harm Falling Person has a duty to bear, if necessary, to prevent her body from killing Victim when she bears no moral responsibility for the threat it poses to him. As I mentioned, Kamm and Frowe believe that the maximum harm that Falling Person might have a duty to bear is greater than that which she might have a duty to bear to prevent someone else's body, or some other threatening object, from killing an innocent person.⁴² This is because they believe that the duty to prevent one's own body from harming an innocent person is stronger than the duty to prevent another body or object from causing equivalent harm to that same person. For Kamm the explanation of this is that if one's body threatens

42 Tadros has subsequently changed his view and now accepts that it is in general impermissible to kill a nonresponsible threatener. His new position is defended in an as-yet-unpublished manuscript called “Why it is Wrong to Kill Non-Responsible Threats.”

the life of another, one is “in an inappropriate relation to the other person.”⁴³ For Frowe, as we have seen, it is that one has a special responsibility for one’s body because one *is* one’s body.

I believe that the claim that we are our bodies, in the sense of being numerically identical to them, is false as a matter of metaphysics. But I have argued for this elsewhere and will not rehearse those arguments here.⁴⁴ I will instead attempt to probe our intuitions about whether the reason one has to prevent one’s own body from causing harm when one is not morally responsible for the threat it poses is stronger than the reason to prevent someone else’s body, for which one is also not responsible, from causing equivalent harm.

Here are four examples intended to elicit intuitions about the comparative importance of preventing one’s own body from killing an innocent person.

Equal Harm A stranger and I have both been captured by Villain, who has already rendered the stranger unconscious and is about to anesthetize me as well. But first he offers me a choice. Either he will drop my unconscious body on Victim 1, thereby killing him, or, if I provide Villain with some benefit at substantial personal cost, he will drop the stranger’s unconscious body on a different innocent person, Victim 2, thereby killing her. There are no relevant differences between Victims 1 and 2 and, as in all such cases, neither the stranger nor I will be harmed.

Lesser Harm Either Villain will drop my unconscious body on Victim 1, causing him harm almost but not quite as bad as the harm involved in being killed, or, if I provide him with the benefit, he will drop the stranger’s unconscious body on Victim 2, thereby killing her.

Brainless Killing 1 I live in a technologically advanced society in which people’s lives can be extended for hundreds of years. In this society, people periodically have their brains temporarily removed from their bodies so that their bodies can be furnished with new organs grown from stem cells. During this process, a person’s brain is able to communicate with other people. I see that some stranger’s body from which the brain has temporarily been removed is about to be dropped by Villain on Victim, with the usual consequences. I can, however, prevent the stranger’s

43 Kamm, “The Insanity Defense,” p. 65.

44 Jeff McMahan, *The Ethics of Killing: Problems at the Margins of Life* (New York: Oxford University Press, 2002), chapter 1; and Tim Campbell and Jeff McMahan, “Animalism and the Varieties of Conjoined Twinning,” in Stephan Blatti and Paul Snowdon, eds., *Animalism: New Essays on Persons, Animals, and Identity* (Oxford: Oxford University Press, 2016). Also see Derek Parfit, “We Are Not Human Beings,” also in Blatti and Snowdon.

body from being dropped if I bear some cost. Suppose that the maximum cost I am morally required to bear to prevent this killing is x .

Brainless Killing 2 My brain has been removed while my body is in the shop for service. I learn that Villain has abducted my body and is about to drop it on Victim. If, however, I agree to bear a certain cost, I can prevent that.

In thinking about these cases, I feel a slight inclination to think that I ought to bear the cost in *Equal Harm* and *Lesser Harm* and that I ought to bear a cost greater than x in *Brainless Killing 2* to prevent my own body from being used to kill Victim. But on reflection I think this is just a product of a certain squeamishness about being causally involved in a killing. It is rather like the feeling of relief that one of my students once told me he had experienced when his family's dog died on the way to the veterinarian's office where she was to be euthanized. Even though the student knew that it would be good for the dog to die, he was deeply averse to being causally involved in her being killed. In much the same way, I might be motivated to ensure that it is the stranger's body that kills an innocent person rather than my own, but not because I believe that I *ought* to. Indeed, I think that in *Lesser Harm* I ought *not* to bear the cost required to induce Villain to drop the stranger's body rather than mine.

In summary, a nonresponsible threatener cannot have a duty not to kill an innocent person involuntarily and nonresponsibly. A fortiori, she cannot have a secondary duty to bear harm in the fulfillment of such a duty; nor can her potential victim or third parties have a permission to harm her as a means or side-effect of enforcing such a duty – that is, one that she does not have. And even if a nonresponsible threatener could acquire a duty by posing a direct threat, it would not derive from her special relation to her body.

I will conclude this section by presenting three further objections to Frowe's six-stage argument for the permissibility of killing direct threateners. The first concerns point 6 above. Frowe accepts that a direct threatener who does not intend the threat she poses is not morally required to sacrifice her life, or accept its sacrifice, to prevent herself from killing Victim. Her enforceable duty is to bear only some lesser, nonlethal harm. Yet Frowe claims that it is still permissible to kill her. This is in part because she thinks that most of the harm the threatener would suffer in being killed is harm that she has a duty to bear. If we then compare the remainder of the harm in being killed – the difference between the amount of harm she has a duty to bear and the full harm of death – we find that it is so much less than the full harm of death for Victim, that Victim or a third party is justified in inflicting it on the ground that it is the lesser evil. Frowe calls the full justification a "hybrid justification" because

it combines two forms of justification for harming: a duty-based justification and a lesser-evil justification.

Suppose, however, that I am right that the harm that a direct threatener (who, again, I am assuming is nonresponsible) has a duty to bear to prevent herself from killing Victim is no greater than that which a nonresponsible *non-threatener*, such as Potential Rescuer, would have a duty to bear in rescuing Victim from a lethal threat. In that case, the hybrid justification can succeed only in cases in which the harm the threatener would suffer in dying is substantially less than that which Victim would suffer, for example because of a great difference in age. Assuming that the amount of harm that a nonresponsible, nonthreatening rescuer has a duty to bear in saving an innocent person is small in relation to the harm of being killed, the difference in magnitude between that harm and the full harm of death is generally too great for its infliction to be justified as the lesser evil when the alternative is *allowing* Victim to be killed.

The second objection is that Frowe's argument has what I believe to be a seriously counterintuitive implication. In another version of the dry well case,

Third Party Intervention Villain has forcibly captured both Victim and Falling Person and rendered them both unconscious. He has placed Victim at the bottom of the well and is about to throw Falling Person down, again with the usual consequences. Because they are both unconscious, neither Victim nor Falling Person can do anything to prevent Victim from being killed. There is, however, an unrelated, disinterested person, Third Party, who has a weapon that he can use, at no cost to himself, to disintegrate Falling Person, though he cannot use it to kill Villain, who is screened from him.

According to Frowe, there is a hybrid justification for killing Falling Person, both elements of which extend to Third Party. Falling Person has an enforceable duty to bear most of the harm of death and Third Party can, at no cost to himself, enforce that duty; therefore he ought to do so. There is also a lesser-evil justification for inflicting the remainder of the harm of death on Falling Person and Frowe believes that agents are required and not merely permitted to act on lesser-evil justifications; hence Third Party ought to inflict the remaining part of the harm of death as well.⁴⁵ That is, he is morally required to kill Falling Person. Yet Third Party has a choice between intentionally killing an innocent,

45 Helen Frowe, "Claim Rights, Duties, and Lesser-Evil Justifications," *Proceedings of the Aristotelian Society Supplementary Volume* 89 (2015): 267–85.

nonresponsible person as a means (albeit an eliminative means) of saving the life of another innocent, nonresponsible person, and unintentionally allowing the second of these people to be killed. Intuitively, and according to widely accepted deontological principles, it seems that Third Party's intervention to kill Falling Person is not only not morally required but is in fact morally prohibited.

The third objection is less serious but worth considering. As I have noted, when attempting to show that it is permissible to kill a direct threatener in defense of a single victim, Frowe claims that it is significant that the threatener "is going to harmfully interfere with an innocent person's body" and that, "since averting such interference requires only non-exploitative, eliminative force, it is permissible for a morally innocent person to use proportionate force to prevent such harmful interference."⁴⁶ But when Victim is about to vaporize Falling Person, he is going to harm an innocent person's body and averting that harm requires only eliminative force. Thus, according to Frowe's view, Falling Person may permissibly kill Victim in counter-defense, if she can.

Frowe will of course respond that the cases are different because Victim is acting permissibly and thus does not have an enforceable duty not to kill Falling Person. It seems, however, that the only reason, on Frowe's view, that Victim's defensive action is permissible while Falling Person's is not is simply that Falling Person threatened *first*. Kamm explicitly says this difference is morally significant. She writes that a nonresponsible threatener "is in a position where she should not be first. ... She should bear a relatively greater burden for correcting this inappropriate position."⁴⁷ It seems to me, however, that this difference in timing is morally insignificant given that Falling Person is no more responsible for threatening first than she is for threatening. Indeed, if there is a morally significant difference, it is that Victim threatens as a morally responsible agent whereas Falling Person does not.

5 Further Intuitive Arguments against the Permissibility of Killing Nonresponsible Threateners

There is another way in which, on Frowe's view, whether it is permissible to kill a nonresponsible person who causes a lethal threat is a matter of timing. Suppose that, in

⁴⁶ Frowe, *Defensive Killing*, p. 9.

⁴⁷ Kamm, *Creation and Abortion*, p. 47.

Reparation 1 Villain throws unconscious Falling Person down a well at the bottom of which lies unconscious Victim. Falling Person's landing on Victim does not kill him but does irreversibly damage one of his vital organs. Her landing on him *will kill* him soon unless he receives an organ transplant. Third Party arrives moments later. He is, of course, a transplant surgeon who quickly determines that Falling Person, who is uninjured but still unconscious, has a tissue type that makes her an ideal source of a replacement organ for Victim, who is also still unconscious. He can perform the surgery, killing Falling Person but saving Victim.

In my view, it is impermissible for Third Party to take Falling Person's vital organ. (It would also be impermissible for Victim to take it if, *per impossibile*, he could perform the transplant himself.) I infer from this that it is impermissible for Third Party (or Victim) to kill Falling Person during her descent. But on Frowe's view, if Third Party had arrived moments earlier, he would have been permitted (indeed, if my second objection above is correct, *required*) to kill Falling Person as she descended. At that point, she was a direct threatener who could have been killed eliminatively. But she is now what I once referred to as an "innocent cause" of a threat, though here it is better to give her the narrower label, "nonresponsible cause."⁴⁸ Although she caused the threat to Victim's life and *will have* killed him if she is not killed now, she is also, in Frowe's taxonomy, an innocent bystander who must be killed exploitatively if Victim is to be saved from the threat she has caused. So, on Frowe's view, although Third Party would have been required to kill Falling Person to save Victim if he had arrived moments earlier, it is now impermissible for him to kill her to save Victim.

I do not deny that many people – probably most – would accept that it would have been permissible for Third Party to kill Falling Person as she descended to prevent her from killing Victim by damaging his vital organ, but not permissible for him to kill her after she has damaged the organ, even though in both instances he would kill her as a means of preventing her from killing Victim. But I do question whether these intuitions are rationally defensible. In this case, the difference between eliminative and exploitative agency is just a matter of timing. If the defensive killing is done just before Falling Person's body causes the immediate threat to Victim's life, it is eliminative; if it is done just after, it is exploitative.

48 McMahan, "Self-Defense and the Problem of the Innocent Attacker," p. 267. Both in this article and in *The Ethics of Killing* (pp. 405–407), I have used the comparison between a nonresponsible cause and a direct threatener to argue that it is normally impermissible to kill the latter. Here I try to take the argument further.

It may be instructive to compare *Reparation 1* with

Reparation 2 Villain is maliciously attempting to kill Victim by driving a car into him. Although Villain's car strikes Victim, it does not kill him but irreparably damages one of his vital organs. The car then swerves into a tree, knocking Villain unconscious. Third Party, who again is a transplant surgeon, can save Victim but only by replacing Victim's damaged organ with Villain's healthy one.⁴⁹

While he was driving toward Victim, Villain was a culpable direct threatener. But when his car struck Victim, he immediately became a bystander. (Because he is culpably responsible for the threat to Victim's life, he is a culpable bystander. He is thus a counterexample to Frowe's claim that "the idea of a morally responsible bystander" is "conceptually incoherent," for "there is ... nothing with respect to the threat to Victim *for which* a bystander can be morally responsible."⁵⁰) It is uncontroversial that while Villain was driving his car toward Victim, he was morally liable to be killed, assuming that killing him was necessary to prevent him from hitting Victim. Thus, if Third Party had been present then and had had the capacity to kill Villain at no cost to himself, he would have been justified in killing him and probably morally required to do so. He has, however, arrived too late to defend Victim eliminatively. But he can still prevent Villain's wrongful action from killing Victim. To me it seems implausible to suppose that Villain can escape liability to be killed just because Third Party has arrived a moment later rather than a moment earlier. Although I share the intuition that it seems worse to kill Villain exploitatively, I think that intuition is mistaken in this case. I believe that, because Villain is culpably responsible for the situation in which either he or Victim must die, it is a

49 I briefly discussed the challenges that a case of this sort raises in "Justice and Liability in Organ Allocation," *Social Research* 74 (2007): 101–124, pp. 111–13, though without at the time appreciating the possible significance of the distinction between eliminative and exploitative agency.

50 Frowe, pp. 29 and 28. Another counterexample is a person who culpably *allows* a threat for which he is neither causally nor morally responsible to eventuate in harm when he ought to prevent that. Frowe argues (pp. 29–30) that such a person is not a counterexample because, although he is culpable, he is not culpable for the existence of the threat and is "not culpable qua bystander." I think that, even if his allowing the threat to continue does not *cause* the threat to continue, he is nevertheless morally culpable for its continuation. And he is a bystander when he allows it to continue, as he would have to be harmed exploitatively to make him fulfill his duty to eliminate the threat. (There is no suggestion in Frowe's example that he might be a moral obstacle rather than a bystander.)

matter of justice that he should bear the burden of his own wrongdoing, even if that requires that he be killed exploitatively.⁵¹

Reparation 1 and *Reparation 2* have the same structure. In each, a person is initially a direct threatener (in one case a nonresponsible direct threatener, in the other a culpable direct threatener) but, having caused a soon-to-be-fatal injury, then becomes a bystander who is also the cause of a present threat (in one case a nonresponsible cause, in the other a culpable cause). In *Reparation 2* I find it difficult to believe that Third Party could be morally required to kill Villain as he drives towards Victim but morally forbidden to kill him a moment later, when in both instances killing him is necessary to prevent his wrongful action from killing Victim. Because I think it clear that Third Party is required to kill Villain eliminatively as he is driving towards Victim, I conclude that he is permitted to kill him exploitatively shortly thereafter.

Because the cases are structurally identical, I think that, in *Reparation 1*, the two ways in which Third Party might kill Falling Person in defense of Victim should also be either both permissible or both impermissible. Because I think it clear that Third Party may not kill Falling Person exploitatively when she has become a nonresponsible cause and thus an innocent bystander, I conclude that he would not have been permitted to kill her eliminatively if he had been present while she was falling. As Falling Person is, during her descent, a paradigm instance of a direct threatener, this suggests that it is normally impermissible to kill such a person in defense of a single victim.

I will conclude by presenting one more pair of examples. Like the preceding argument, and like Otsuka's original argument and Frowe's argument that appeals to cases involving a flagpole, a shield, and so on, it takes the form of a comparison between a case involving a direct threatener and another case that is intended to be morally indistinguishable. These final two examples together challenge the moral significance of one of the considerations that Frowe thinks is particularly important – namely the distinction between direct and indirect threateners.

Duress 1 Villain captures and sedates an innocent person, Unconscious Victim, whose death he desires. He also captures Victim, whom he wants to coerce to kill an innocent person. He puts a gun to Victim's head and orders him to kill Unconscious Victim. Victim believes, correctly, that it is certain that he will be killed if he refuses. He cannot use the means he has for killing Unconscious Victim to kill Threatener instead.

51 Frowe (p. 163) calls Third Party's killing of Villain "opportunistic" rather than "exploitative," to mark the fact that Villain is responsible for the threat to Victim.

Next suppose that these same people are the dramatis personae in a slight variant of Frowe's *Ray Gun*.

Duress 2 Villain wants to coerce Victim to kill an innocent person. He places Victim at the bottom of a narrow, dry well and throws him a weapon capable of disintegrating a human body. He then drops Unconscious Victim's unconscious body down the well. Victim can survive only by disintegrating Unconscious Victim.

The common and I believe correct understanding of *Duress 1* is that it is impermissible for Victim to kill Unconscious Victim. If he does kill her, he may not be blameworthy but he will have acted wrongly. The common intuition about *Duress 2*, however, is that it is permissible for Victim to kill Unconscious Victim in self-defense. Frowe, I think, would argue that the relevant difference is that in *Duress 1* Unconscious Victim is an innocent bystander who must be killed exploitatively, whereas in *Duress 2* she is a direct threatener whose killing would be eliminative. In support of the view that Unconscious Victim is a bystander in *Duress 1*, one might claim that Victim requires her to be present so that he can kill her as a means of obeying Villain's command.

But perhaps this is a mistake. Unconscious Victim in *Duress 1* may not be an innocent bystander. What Victim actually needs is for Unconscious Victim not to be present, both earlier and at present. Had Unconscious Victim not been present at all, Villain would have had no occasion to coerce Victim by threatening his life. And if Unconscious Victim were miraculously to vanish, the threat to Victim would vanish with her, as Villain would then have no reason to kill him. It seems, in other words, that Unconscious Victim is, in Frowe's taxonomy, an indirect threatener, as her physical presence is necessary for there to be any threat to Victim. (Categorization is problematic here because Unconscious Victim is not a physical obstructor. Perhaps she is an unusual instance of a moral obstacle who must be killed eliminatively, to eliminate her contribution to the threat, rather than as a side-effect.)

If this is right, then in both *Duress 1* and *Duress 2*, Victim's killing of Unconscious Victim would be eliminative rather than exploitative. In both cases, the person who is both causally and morally responsible for the threat to Victim is Villain. In both cases, to defend himself successfully, Victim must kill a morally innocent person who has not exercised any form of agency in contributing to the threat that he faces. The only difference is that in *Duress 2*, Unconscious Victim's body is the instrument with which Villain threatens Victim's life, whereas in *Duress 1* it is a gun.

I cannot detect any morally significant difference between the two killings. If I am right that there is no significant difference, these cases support the second premise in the restatement of Otsuka's original argument in Section 2. Because I suspect that the intuition that it is permissible for Victim to kill Unconscious Victim in *Duress 2* is an overgeneralization of intuitions about standard cases of self-defense in which it is permissible to kill someone who threatens one's life, I think my intuition that it is impermissible for Victim to kill Unconscious Victim in *Duress 1* is more reliable. I therefore conclude that it is impermissible to kill Unconscious Victim in *Duress 2*. Because Unconscious Victim in *Duress 2* is just Falling Person in an unconscious state, and Falling Person is, as I noted earlier in this section, a paradigm instance of a direct threatener, I think we should conclude – again – that it is normally impermissible to kill such a person in defense of a single victim.

Biographical Note

Jeff McMahan (born 30 August 1954) is White's Professor of Moral Philosophy at the University of Oxford, and taught previously at Rutgers University (2003–2014) and the University of Illinois, Urbana-Champaign (1986–2003).