

## The Moral Status of Nonresponsible Threats

---

JASON HANNA

**ABSTRACT** *Most people believe that it is permissible to kill a nonresponsible threat, or someone who threatens one's life without exercising agency. Defenders of this view must show that there is a morally relevant difference between nonresponsible threats and innocent bystanders. Some philosophers, including Jonathan Quong and Helen Frowe, have attempted to do this by arguing that one who kills a bystander takes advantage of another person, while one who kills a threat does not. In this paper, I show that the proposals offered by Quong and Frowe have unacceptable implications. I then argue that those who claim that nonresponsible threats may be killed face a dilemma generated by the possibility of a stationary threat, or someone who endangers another person's life without moving. Unless we arbitrarily distinguish between stationary and moving nonresponsible threats, it is unclear how the permission to kill nonresponsible threats is to be explicated. I conclude that nonresponsible threats are not legitimate targets of self-defence.*

### 1. Introduction

Is it permissible to kill a nonresponsible threat? As I shall understand it here, a nonresponsible threat is someone who endangers another person's life without thereby exercising agency.<sup>1</sup> Most people accept what I shall call the 'traditional view', which would permit you (though perhaps not a third party) to kill in the following case:

*Fall:* An evil-doer has thrown Falling Person down a well. Through no fault of your own, you are standing at the bottom of the well. Falling Person will crush you to death if you do not first disintegrate his body with your ray gun. If he does land on you, however, he will survive. You kill Falling Person.<sup>2</sup>

The standard objection to the traditional view is that nonresponsible threats such as Falling Person are relevantly like bystanders. For present purposes, we can understand a bystander to be someone who is not causally involved in, and is not morally responsible for, the threat to one's life.<sup>3</sup> Since it is clearly impermissible to kill bystanders, the objection goes, it is likewise impermissible to kill nonresponsible threats.<sup>4</sup> Recently, some defenders of the traditional view, including Jonathan Quong and Helen Frowe, have challenged this analogy by arguing that the killing of bystanders involves a more objectionable form of agency than the killing of threats. This line of argument holds that one who kills a bystander, unlike one who kills a threat, *takes advantage* of another person.<sup>5</sup>

I have two goals in this article. First, I argue against this defence of the traditional view. Even if it points to a relevant difference between nonresponsible threats and some bystanders, it fails to distinguish between nonresponsible threats and others who cannot permissibly be killed. I show (in Sections 3 and 4, respectively) that Quong and Frowe fail to offer a convincing response to this problem. Second, I offer a new argument

against the permissibility of killing nonresponsible threats. This argument appeals to cases in which a *stationary* person nonresponsibly threatens another's life. According to this argument, either the traditional view draws morally arbitrary distinctions or it is unclear how that view is to be explicated. I conclude that the prospects for the traditional view are bleak.

Although cases involving nonresponsible threats may be very rare, the issue I shall discuss merits serious attention by those hoping to develop a theory of self-defence. Some philosophers have argued that one is permitted to kill in defence of one's own life *only* if the other person is culpable or at least morally responsible for endangering one's life, so that the relevant notion of culpability or responsibility justifies self-defence in every case in which it is justified.<sup>6</sup> If it is permissible to kill nonresponsible threats, however, these views must obviously be rejected: although a threat's responsibility might still be relevant (for instance, in determining whether the threat may employ counter-defence), we would need to identify some further feature sufficient to justify a prospective victim's use of defensive force in nonresponsible threat cases.<sup>7</sup> A complete theory of self-defence must therefore address the problem of nonresponsible threats.<sup>8</sup>

In the following, I will consider cases in which a nonresponsible threat endangers just one person. I will set aside cases involving multiple victims or multiple threats, since such cases may raise distinct moral problems. Further, since some proponents of the traditional view claim that a third party may not kill a nonresponsible threat to another person's life, I will focus on cases in which the threatened party is able to defend herself without assistance from others.<sup>9</sup>

## 2. Nonresponsible Threats and Opportunistic Agency

It may seem intuitively clear that you are permitted to kill the threat in *Fall*. Nonetheless, as several philosophers have argued, there seems to be a strong reason for thinking that *Fall* is relevantly like the following case:

*Shield*: An aggressor is shooting at you. You grab a bystander, Shield, and position her in front of you in order to shield yourself from the bullets. Had you not done so, Shield would have survived.

In both *Fall* and *Shield*, you kill an innocent person when doing so is necessary to save your life. Although Falling Person's body may be involved in the threat to your life, his agency is not, and in this respect he is like Shield. Indeed, Michael Otsuka goes so far as to suggest that a nonresponsible threat such as Falling Person can be regarded as a bystander to his body, over which he has no control.<sup>10</sup> Virtually everyone agrees that you act impermissibly in killing Shield. It may then seem that your act of killing Falling Person is likewise prohibited.

As Quong and Frowe (separately) argue, however, there is a potentially important difference between the two cases: you profit from the presence of Shield, but not Falling Person. If Shield's body were not available for your use, you would be unable to save your life. Borrowing a term from Warren Quinn, we might describe your harmful agency in this case as *opportunistic*, since Shield's presence offers an opportunity that you exploit.<sup>11</sup> You do not similarly exploit Falling Person. Although your survival depends on your killing Falling Person, it obviously does not depend on Falling Person's presence, since

he poses the problem that you are trying to eliminate. Again following Quinn, we might describe the harmful agency in this case as merely *eliminative*.<sup>12</sup> Opportunistic agency seems more objectionable than eliminative agency insofar as it displays a greater attitude of disrespect towards the victim. The agent who acts opportunistically is ‘grateful’ that another person is there for him to use.<sup>13</sup> To be sure, the alleged asymmetry between opportunistic and eliminative agency does not by itself show that it is permissible to kill nonresponsible threats. Instead, the asymmetry points to a flaw in the analogy between nonresponsible threats and bystanders. If this analogy is mistaken, however, the most serious objection to the traditional view can be dispensed with. Call this response ‘the opportunistic agency argument’.

I shall grant for the moment that the opportunistic agency argument does point to a morally relevant difference between *Fall* and *Shield*. As Quong recognizes, however, the argument does not seem to account for the intuitive wrongness of some self-perservative killings. Consider an innocent obstructor — someone who non-culpably blocks a threatened party’s only path to safety:

*Alcove:* You are in a tunnel and can move safely out of the way of an approaching train only by taking refuge in an alcove. Unfortunately, Obstructor is already in the alcove, which can accommodate only one person. To save your life, you disintegrate Obstructor’s body with your ray gun and enter the alcove.<sup>14</sup>

In this case, you clearly do not profit from Obstructor’s presence: if he were absent, you could take shelter in the alcove and save your life. Your killing thus seems eliminative, rather than opportunistic. Nonetheless, it seems impermissible to kill Obstructor. And if it is impermissible to kill Obstructor even though doing so involves only eliminative agency, the opportunistic agency argument is in trouble. Although the argument may reveal a difference between *Fall* and *Shield*, it apparently does not reveal a difference between *Fall* and *Alcove*.<sup>15</sup>

There are two responses to this problem available to defenders of the traditional view. They could argue either that obstructor killings typically do involve opportunistic agency, or else that there is some other relevant difference between obstructors and nonresponsible threats. Quong adopts the first of these strategies, while Frowe pursues the second. I consider each in turn.

### 3. Opportunistic Agency and Rights Over Space

Quong offers a two-step argument that your killing of Obstructor would be opportunistic. First, he claims that one can opportunistically use a person, in the sense that we find especially objectionable, by exploiting something that belongs to that person or to which that person has a right. For instance — my example — if a deep-sea diver fends off a shark attack by stealing a fellow diver’s life-sustaining oxygen tank and using *it* as a shield, he has used or exploited the fellow diver. Indeed, this case seems morally on a par with *Shield*, above. Second, Quong argues that people have moral claims not only to their bodies, but also to the space occupied by their bodies. For example, it seems that you would infringe a person’s rights if you were to move her body, even if it were possible to do so without injuring her or even touching her.<sup>16</sup> Quong concludes that it is permissible to use necessary and proportionate force in order to ‘defend something that

belongs to you' if and only if 'you would be able to keep what belongs to you if the other person *and all their property* [including the space they occupy] were not present'.<sup>17</sup> Call this 'the counterfactual test'.

According to Quong, in order to determine whether you would be able to save your life if the other person's space were not present, we must ask whether you would be able to save your life if the other person's space were 'inaccessible to you (e.g., protected by a barrier)'.<sup>18</sup> This proposal initially appears to deliver the desired results. In *Alcove*, you can save your life only by occupying space that is currently occupied by Obstructor; if the space inside the alcove were blocked by a physical barrier, you would have no way to escape the runaway train. The counterfactual test thus implies that you may not kill Obstructor. In *Fall*, by contrast, you do not need to use Falling Person or his space, and the counterfactual test thus permits your defensive killing.<sup>19</sup>

But Quong's counterfactual test has counter-intuitive implications in other cases, as is revealed by the following example:

*Lethal Gas*: You suddenly become very ill and can save your life only by manufacturing a certain drug. Fortunately, you have the necessary ingredients ready to hand. Unfortunately, the manufacture of the drug would release dangerous fumes. These fumes would be blown away from you and through your neighbour's open window before reaching lethal concentrations in his house. Although you know that your neighbour is at home, there is insufficient time for you to warn him, and you are unable to move in order to ensure that the fumes are released in a safer location. You manufacture the drug, and your neighbour dies.<sup>20</sup>

Now in this case it is obvious that you do not profit from the presence of your neighbour's body. To determine whether you profit from your neighbour's space, Quong would have us ask whether you would be able to save your life if this space were inaccessible. So imagine that your neighbour's property — the space he occupies, his house, his yard, and so on — is blocked on all sides by a barrier, such as an infinitely high brick wall. Now suppose that the wind is such that, if this barrier were present, the dangerous gas would be blown away from you and down a side of the barrier before dispersing to safe levels. Under these circumstances, you would be able to save your life without using your neighbour's space; by hypothesis, if that space were inaccessible, the dangerous fumes would be harmlessly blown somewhere else. While you may be able to save your life in the actual circumstances only by sending a harmful substance into space that belongs to your neighbour, you do not *profit* from the presence or availability of this space: you would be no worse off if this space were completely obstructed. Quong's counterfactual test thus seems to imply that you may manufacture the drug, even though doing so will result in your neighbour's death. Yet such behaviour is intuitively impermissible. Your neighbour, after all, is a bystander to the threat you face.

This implication seems sufficiently objectionable to merit the rejection of the counterfactual test. Of course, even if I am correct that the counterfactual test fails to distinguish permissible from impermissible killing, it may be that behaviour which fails the test is, all else equal, more objectionable than behaviour which passes the test. But there are grounds for rejecting even this weaker view. To see why, let us alter one detail of *Lethal Gas*. Suppose that the wind has changed, so that if your neighbour's property were blocked by an impenetrable barrier, the fumes would not disperse very well and

would instead reach a lethal concentration near you. In the actual circumstances, however, the wind will still carry the fumes through your neighbour's open window. In this version of the example, you are unable to save your life without using space that belongs to your neighbour. If that space were inaccessible, so that your neighbour's open window could not be used as an outlet for the dangerous fumes, then you would be killed by the dangerous fumes. The counterfactual test thus condemns your self-preservative behaviour in this case. Yet it is very difficult to believe that there is *any* morally relevant difference between the first variant and the second variant of *Lethal Gas*. Something as trivial as the speed of the wind would not seem to make your behaviour any more or less objectionable.

So far as I can tell, there is just one way for defenders of the counterfactual test to avoid the problems I have raised. This would be to claim that your neighbour has rights over the space in which you manufacture the drug. If this space were inaccessible to you, you would be unable to save your life in both versions of *Lethal Gas*. (Recall that we are assuming that you cannot move to a safer location and are able to manufacture the drug *only* in a space from which the fumes will be blown into your neighbour's house.) Perhaps a critic would argue that your neighbour has rights over this space, just as one might plausibly argue that an individual has rights over the air surrounding his body on the grounds that it will be breathed into his body.<sup>21</sup> This strategy seems unlikely to succeed, however. For the purposes of the example, we can suppose that you manufacture the drug on your own property — in your yard or garage, for instance — before the wind carries the fumes into your neighbour's house. Under these circumstances, you have a prior claim over the space in which you manufacture the drug.<sup>22</sup> Nonetheless, your behaviour is still intuitively impermissible. Of course, we might account for the wrongness of your behaviour by claiming that you have violated your neighbour's right not to be assaulted with lethal fumes. But Quong cannot argue that the fact that you have violated this right shows that you have exploited space that belongs to your neighbour, since Quong's argument implies that whether you have violated (rather than permissibly infringed) your neighbour's rights *depends* on whether you have exploited his space.<sup>23</sup>

I have argued that the counterfactual test permits the killing of some bystanders and draws arbitrary distinctions among cases. Even if we set aside the problems discussed so far, however, Quong's counterfactual test seems to face further difficulties in a set of cases we have not yet considered. Imagine two agents who compete for the use of a life-saving resource that neither of them owns, as in the following example:

*Alcove (Competition)*: Both you and Speedy are (innocently) located inside a train tunnel. In order to escape an approaching train, each of you runs from opposite directions towards an alcove large enough to accommodate just one person. It soon becomes apparent that Speedy will get there first. In order to save your life, you shoot and kill Speedy before he reaches the alcove.<sup>24</sup>

In this case, unlike the original *Alcove* case, you do not profit from the availability of space occupied by the other person. Speedy does not occupy the space in the alcove, and you do not need to pass through any of his space in order to reach the alcove. It thus appears that Quong should classify your agency here as eliminative rather than opportunistic. And since Quong's view distinguishes between the killing of bystanders and the killing of nonresponsible threats solely by appeal to the distinction between eliminative and opportunistic agency, his view will fail to distinguish your killing of Speedy in *Alcove*

(*Competition*) from your killing of Falling Person in *Fall*. Nonetheless, it seems clear that you have acted impermissibly in killing Speedy: although you may race Speedy to the alcove in a fair competition, you may not shoot him in order to ensure your victory. If this is correct, we have further reason to reject Quong's counterfactual test.

#### 4. Direct and Indirect Threats

Return now to the problem that the counterfactual test was intended to solve: even if the opportunistic agency argument distinguishes nonresponsible threats from some bystanders, it does not distinguish nonresponsible threats from innocent obstructors. A second response to this problem is offered by Helen Frowe. Frowe, unlike Quong, would argue that obstructor killings involve only eliminative agency. This is because on her view obstructors are properly classified as threats. Frowe's argument for this claim rests on a more general account of threats. According to Frowe, 'a threat is anyone who makes Victim worse off for her movements or presence'.<sup>25</sup> To be clear, a threat may not *actually* leave the victim worse off. In *Fall*, Falling Person will not leave you worse off if you vaporize him before he lands on you. But if you refrain from killing Falling Person, you will be worse off than you would have been if he had not been present. Likewise, if you refrain from killing Obstructor in *Alcove*, you will be worse off than you would have been if he had not been present.<sup>26</sup> For this reason, Obstructor is a threat on Frowe's account.

This does not mean that Obstructor can be permissibly killed, however. Frowe instead argues that there is a morally relevant distinction between 'direct threats' and 'indirect threats'. A direct threat is someone who will kill the victim if the victim does not take defensive measures. An indirect threat makes the victim worse off in the sense described above but does not threaten to kill the victim. To illustrate, Frowe argues that a 'typical example of an indirect threat is Mafia Boss, who hires Assassin to kill Victim'; Assassin is a direct threat, since it is he, not Mafia Boss, who will actually kill Victim.<sup>27</sup> On Frowe's view, an (innocent) agent may kill a *direct* threat to her life regardless of the threat's culpability, while an (innocent) agent may kill an *indirect* threat to her life only if the threat is culpable — that is, if the threat intentionally passed up an 'opportunity to avoid exposing Victim to harm'.<sup>28</sup> Since Obstructor will not *kill* you if you refrain from killing him, this view implies that he is an innocent indirect threat who cannot permissibly be killed.

Frowe is not alone in claiming that the important difference between Falling Person and Obstructor is that only the former threatens to kill. Judith Jarvis Thomson, for example, argues that it is permissible for you to kill Falling Person because he will otherwise violate your rights, and that he will otherwise violate your rights because he will otherwise kill you.<sup>29</sup> By contrast, Thomson claims, innocent obstructors do not threaten to kill anyone.<sup>30</sup> Suzanne Uniacke argues that the right of self-defence (and other-defence) is a 'corollary of the victim's right not to be *killed*'.<sup>31</sup> In short, many prominent defenders of the traditional view seem to believe that the important difference between nonresponsible (direct) threats and obstructors is that only the former will kill their victims.

For Frowe's view to be successful, we need some reason for thinking that the distinction between direct and indirect threats is morally relevant. Yet it is far from clear that such a reason can be offered. Indeed, a comparison of several cases seems to show that



the distinction does not have the significance that Frowe attributes to it. Recall Frowe's claim that Mafia Boss is an indirect threat, since he will not kill Victim. In the following case, a nonresponsible person plays a causal role equivalent to that of Mafia Boss:

*Look-Alike:* You know that the mob has developed a code according to which a mafia boss, by making a particular hand gesture, can order the assassination of the person sitting nearest to him. You are sharing a public bench with Look-Alike Lou, who (unbeknownst to him) bears an uncanny resemblance to a well-known mafia boss. As a result of an uncontrollable muscle spasm, Lou is now making the dreaded hand gesture. You notice that a well-known mafia assassin is about to turn in your direction. There is no time for you (non-lethally) to prevent Lou from making the gesture, and no way for you to kill the assassin. You can save yourself only by disintegrating Lou's body with your ray gun. You do so.

If Mafia Boss is an indirect threat in Frowe's original example, then surely Lou is an indirect threat in this case; in both cases, after all, it is the assassin who threatens to kill you. Since Lou is a non-culpable indirect threat, Frowe's view implies that you may not kill him.

But now consider a variant of *Look-Alike*: suppose that the role of the assassin is played not by a human agent, but instead by a sophisticated robot capable of recognizing the mob's coded hand gestures. In this case, where no agent will intervene between Lou's hand gesture and your death, it seems that Lou is a direct threat. Notice that it would be implausible for Frowe to challenge this claim by arguing that the robot, rather than Lou, will kill you. After all, if Lou's movement were about to result in his firing a gun or setting off a motion-sensitive bomb, Frowe surely would not want to claim that Lou is an indirect threat on the grounds that it is the bullet or bomb that will kill you; these cases are precisely the sort in which the traditional view is designed to permit defensive killing. And if Lou is a direct threat when his uncontrollable movement will activate a lethal bomb, surely he is also a direct threat when his uncontrollable movement will activate the lethal robot.

Although Frowe should apparently classify Lou as an indirect threat in the human assassin case and a direct threat in the robot assassin case, this difference does not appear to make a moral difference. I find it difficult to believe that you are morally permitted to kill Lou if his gesture will initiate the activity of a non-responsible robot, while you are morally forbidden from killing Lou if his gesture will initiate the activity of a responsible human assassin. Unless we are prepared to accept this result, we must apparently reject the view that there is a morally relevant distinction between direct threats and indirect threats. For those who do not share my intuition that there is no morally relevant difference between the two versions of *Look-Alike* considered so far, however, it is worth briefly considering one further variant: suppose that the role of the assassin is played by an innocent person who (as you know) has been hypnotized, against his will, so that he will respond to the mob's hand gesture. Frowe would presumably classify the hypnotized assassin as a direct threat, since she would hold that you may kill him. And if the hypnotized assassin is a direct threat, then Frowe should classify Lou as an indirect threat, for the same reason that she classifies Mafia Boss as an indirect threat when he orders Assassin to kill. Yet defenders of the traditional view would presumably agree that it is arbitrary to claim that you may kill Lou if (as a direct threat) he will initiate the

activity of the robot but not if (as an indirect threat) he will initiate the activity of the hypnotized assassin. Frowe's view seems plausible insofar as it fits our intuitions about cases: while it is intuitively permissible to kill direct threats such as Falling Person, it is intuitively impermissible to kill indirect threats such as Obstructor. As we have now seen, however, this proposal yields other implications that are intuitively unacceptable.

I have considered two attempts to show that there is a morally relevant difference between nonresponsible (direct) threats and obstructors. Both have implications that defenders of the traditional view should find problematic. Of course, even if I have been correct so far, it may be thought that defenders of the traditional view have available some further reply that I have not yet considered. I turn now to an argument that seems to challenge the traditional view more generally.

## 5. Stationary Nonresponsible Threats

This final argument appeals to a character we have not yet encountered and that is perhaps best described as a *stationary nonresponsible threat*. (Although I shall refer to such a person as a 'threat', I do not necessarily mean to suggest that the traditional view would permit him to be killed.) A stationary nonresponsible threat jeopardizes another person's life without exercising agency or even moving. Here is an example:

*Fall (Inverse)*: An evil-doer has pushed you off the top of a cliff overlooking the sea. You are an accomplished diver and would ordinarily be able safely to enter the water. Unfortunately, however, Stationary Person is directly in your trajectory, lying asleep and completely still on an inflatable mattress. If you collide with Stationary Person, you will be killed by his exceptionally bony knees, though he will survive unharmed. There is no way for you to wake Stationary Person in time. You can save your life only by disintegrating his body with your ray gun.

Stationary Person is in one respect like Obstructor: he blocks your path to safety. Yet Stationary Person is in another respect like Falling Person: if you do not kill him, the collision with his body will result in your death. It may not be immediately clear what defenders of the traditional view should say about the permissibility of your killing Stationary Person. In the following, I shall argue that both possible verdicts raise serious problems for the proposals considered in the previous sections and for the traditional view more generally.

### 5.1. If Stationary Nonresponsible Threats May Not Be Killed

Suppose first that you are *not* permitted to kill Stationary Person. It seems to me, at least, that this is the intuitively correct result. Further, many defenders of the traditional view have good reason to accept it. As we have seen, some defences of the traditional view imply that it is permissible for a (non-culpable) agent to kill an innocent person in order to save his own life only if that person will otherwise kill him. And it does not appear that Stationary Person will *kill* you if you refrain from killing him. We might accept Thomson's claim 'that if an event that consists in the fall of Y on X kills X, then it follows that Y killed X'.<sup>32</sup> But it seems more difficult to accept the inverse claim: that if an event that



consists in the fall of Y on X kills Y, then it follows that X killed Y. Analogously, we would not say that you have punched someone who quickly moves his face into contact with your stationary hand, or that you have assaulted someone who lashes out during a seizure and breaks his hand on your bony knees.

Indeed, it is far from clear that Stationary Person would be responsible for *killing* you even if he were able to control his behaviour. Suppose, for example, that Stationary Person is not asleep when you are pushed over the cliff but instead chooses to stay put for some compelling reason. (Imagine that by moving, he would expose himself to potentially lethal ocean currents.) Under these circumstances, it may be implausible to say that, unless you kill him first, Stationary Person will kill you. After all, he does not take any action that will result in your death. Instead, like an innocent obstructor, he simply refrains from moving out of your way.<sup>33</sup> If Stationary Person's behaviour would not constitute killing even if it were under his control, we have strong reason to conclude that Stationary Person does not threaten to kill you in the original example, in which he is unconscious.

According to defences of the traditional view that permit you to kill an innocent person only if that person will otherwise kill you, it thus appears that you are prohibited from killing Stationary Person in self-defence. It is difficult to reconcile this implication with the judgment that you are permitted to kill Falling Person in *Fall*, however. In both *Fall* and *Fall (Inverse)*, you are made worse off by the presence of the other party. In both cases, defensive killing prevents a state of affairs in which you die as a result of a collision with the other person's body. There is just one difference between the two cases: in *Fall* the threat is moving while the prospective victim is stationary, and in *Fall (Inverse)* the threat is stationary while the prospective victim is moving. Yet it is not at all clear why this difference would matter. That Falling Person is moving may show that Falling Person, unlike Stationary Person, threatens to kill you. As many philosophers have recognized, however, the killing/letting die distinction has a strong claim to moral significance only insofar as it represents a distinction between two forms of *agency*, and thus it does not appear applicable to nonresponsible threats.<sup>34</sup>

In response to my claim that *Fall* and *Fall (Inverse)* are morally on a par, a critic might argue that Quong's counterfactual test identifies an important difference between these cases. In *Fall (Inverse)*, it appears that you are able to save your life only by using space occupied by Stationary Person; if this space were blocked by a physical barrier, you would be unable safely to enter the water. Although I have already criticized Quong's counterfactual test, it is worth pointing out that his view yields odd results when applied to *Fall (Inverse)*. Imagine two different ways in which this case could proceed. In the first, you disintegrate Stationary Person with your ray gun and then enter the water by moving through the space previously occupied by his body. Quong's view does seem to condemn your behaviour here. In the second version of the example, imagine that the recoil of your ray gun is strong enough to alter the trajectory of your fall. After firing at Stationary Person, you enter the water a short distance from the space he occupied. (Assume that you were able to alter your trajectory in the appropriate way *only* by firing in the direction of Stationary Person, so that your action was necessary in order to save your life.) Quong's view seems to permit your behaviour in this second version of the case: if the space occupied by Stationary Person had been blocked by a barrier, you still would have been able to save your life by using the recoil of your gun to direct yourself away from the barrier. Nonetheless, it seems clear that your conduct in the second version of the

case has the same moral status as your conduct in the first. It is rather difficult to believe that whether it is permissible for you to kill Stationary Person depends on whether the recoil of the gun is strong enough to alter your trajectory. Quong's version of the opportunistic agency argument does not seem to provide a satisfying resolution of *Fall (Inverse)*.

The first horn of a dilemma for the traditional view can now be summarized. If it is impermissible for you to kill Stationary Person in *Fall (Inverse)*, it would likewise appear impermissible for you to kill Falling Person in *Fall*. There does not appear to be a morally relevant difference between these cases.

### 5.2. *If Stationary Nonresponsible Threats May Be Killed*

Consider now a second possibility. Suppose that defenders of the traditional view are content to claim that you are permitted to kill Stationary Person. I find this result counter-intuitive. But even if others are willing to accept it, the traditional view faces a new problem: how can the permission to kill nonresponsible threats be interpreted so that it yields the appropriate implications? As we have seen, some philosophers seem to think that the permission to kill such threats follows from the permission to kill someone who will otherwise kill you (assuming that you are not culpable). As we have also seen, however, it is doubtful that Stationary Person threatens to kill you. The defender of the traditional view thus needs to explain what separates Stationary Person from those who you may not kill in order to save your life.

It may be tempting here to appeal to Stationary Person's causal role in the threat to your life. Even if it is odd to say that Stationary Person will *kill* you, one might argue that he will nonetheless *cause* your death unless you kill him first.<sup>35</sup> The relevance of this causal claim may be supported by the way in which 'threat' and 'bystander' are commonly defined. Quong, among others, considers a bystander to be someone who has 'no significant causal involvement in whatever it is that threatens [the victim's] life and [has] done nothing that would normally make [him] liable to be killed'.<sup>36</sup> This definition suggests that someone who does have 'significant causal involvement in whatever it is that threatens' the victim is thereby a threat. Perhaps a critic would argue that you are permitted to kill Stationary Person for the same reason that you are permitted to kill Falling Person: each is a threat insofar as each will otherwise cause your death.

An initial problem for this appeal to Stationary Person's causal role is that it may permit too much killing. Consider *Alcove (Competition)*, described earlier, in which you and Speedy are racing to an unoccupied alcove in order to escape an oncoming train. (Recall that in this case, you need not use Speedy's space, since you can shoot him before he reaches the alcove.) It would certainly appear that Speedy will have 'significant causal involvement' in your death unless you shoot him first. Indeed, your life would not be in danger if Speedy were not present, or if he were present but were not now running towards the alcove. If the proponent of the traditional view hopes to accommodate common-sense moral intuitions, she needs to show that the line of argument under consideration does not permit you to shoot and kill Speedy.

Perhaps a critic would argue that even if Speedy will be in some way *involved* in your death unless you kill him first, he does not threaten to be a *cause* of your death.<sup>37</sup> In contrast, it may appear that Stationary Person does threaten to be such a cause. This response faces several difficulties. First, it is not immediately obvious that the relevant

causal claims could be adequately defended. Some accounts of causation imply that Speedy will be a cause of your death; defenders of the traditional view would thus need to argue against these accounts.<sup>38</sup> Further, although I obviously cannot provide a full-blown account of causation here, it is worth noting that there are important similarities between Stationary Person and Speedy, even beyond their lack of culpability. Neither Stationary Person nor Speedy initiates the sequence of events that threatens your life.<sup>39</sup> Speedy's behaviour ensures that you will be killed by something else, namely the oncoming train or whatever it was that started the train in motion. Perhaps this is why some would argue that Speedy is not a cause of your death. Yet the behaviour of Stationary Person's body likewise ensures that you will be killed (or will have been killed) by something else, namely the evil-doer who initiated the threat to your life by pushing you over the cliff. It is thus unclear that we have any reason for classifying Stationary Person, but not Speedy, as a cause of your death.

Moreover, even if it is true that (assuming you do not kill him first) Speedy will be involved in your death without being a cause of your death, the response we are now considering is in need of further defence. The response will succeed only if the critic can show that there is a morally relevant difference between its being the case that an innocent person will be a cause of your death unless you kill him, on the one hand, and its being the case that an innocent person will be in some other way involved in your death unless you kill him, on the other. Even if there is a difference here, however, it is not at all clear why it should be thought morally relevant.<sup>40</sup> The problems noted so far point to an underlying difficulty facing any attempt to argue that Stationary Person may be killed in virtue of his causal role: it would be unfortunate for proponents of the traditional view if their position rests on contentious metaphysical claims that few of them have sought to defend. If they argue that it is permissible to kill Stationary Person because he will otherwise be a cause of your death, however, they may be forced to confront, and attribute moral importance to, difficult questions about causation.

We are now in a position to see the dilemma for the traditional view. As I argued in Section 5.1, it is difficult to believe that there are any morally relevant differences between *Fall* and *Fall (Inverse)*. Thus, if you may not kill Stationary Person, it would seem that you may not kill Falling Person, either. If, on the other hand, defenders of the traditional view claim that you may kill Stationary Person, they must explain what distinguishes him from others who cannot be permissibly killed in order to save one's own life. As I have argued in this section, it is not at all clear that such an explanation can be given.

## 6. Conclusion

I have offered two arguments against the permissibility of killing nonresponsible threats. First, I rebutted several attempts to show that there is a morally relevant difference between nonresponsible threats and innocent obstructors. Quong's response to this problem appeals to the obstructor's rights over space. This response seems overly permissive and arbitrarily distinguishes among cases. Frowe responds to the threat/obstructor analogy by claiming that an obstructor is an indirect threat and that indirect threats may be killed only if they are culpable. This view likewise draws problematic distinctions, as was evident in the robot and human assassin cases. The analogy between nonresponsible (direct) threats and obstructors thus appears sound. If there still seems

to be an important difference between them, it may be because we think that threats such as Falling Person will be more closely involved in the deaths of their victims. As I showed in the previous section, however, it is possible for a threat to be stationary. This possibility gives rise to a dilemma for the traditional view: either the traditional view draws arbitrary distinctions between cases such as *Fall* and *Fall (Inverse)*, or else it is unclear how the permission to kill nonresponsible threats is to be explicated.

I conclude that the traditional view should be rejected. Although this conclusion may pose a problem for some accounts of self-defence, it does not by itself support any particular account, and I have not here attempted to develop a comprehensive view of the limits or justification of our defensive privileges. Nonetheless, the intuitive appeal of the traditional view has often been thought to pose an objection to theories that ground defensive privileges solely in the culpability or moral responsibility of the threat or aggressor. If I am correct, however, this objection can be overcome.

Jason Hanna, Department of Philosophy, Zulauf Hall 915, Northern Illinois University, DeKalb, IL 60115, USA. [jhanna@niu.edu](mailto:jhanna@niu.edu).

## Acknowledgements

Earlier versions of this paper were presented at the 2011 Rocky Mountain Ethics Congress, University of Colorado, and at the 2010 Joint Meeting of the Illinois and Indiana Philosophical Associations, Eastern Illinois University. I am grateful to those audiences, and especially to Kai Draper and Robyn Gaier, for helpful feedback. I am also indebted to an anonymous referee for the *Journal of Applied Philosophy*.

## NOTES

- 1 Some commentators use the term 'innocent threat' to describe what I am calling a 'nonresponsible threat'. Since innocence contrasts with culpability, however, 'innocent threat' is perhaps best reserved to describe someone who exercises agency but is not culpable for the threat he poses. See Jeff McMahan, *Killing in War* (Oxford: Oxford University Press, 2009), p. 166; and Phillip Montague, 'Self-defence and innocence: Aggressors and active threats', *Utilitas* 12,1 (2000): 62–78, p. 65.
- 2 This example is originally from Robert Nozick, *Anarchy, State, and Utopia* (New York: Basic Books, 1974), pp. 34–5.
- 3 As I shall suggest below (Section 5.2), there may be some people who do not obviously fit this description even though they are commonly regarded as bystanders. The important point here, however, is that most people agree that it is impermissible to kill those who do fit this description, even when such killing is necessary to save one's own life.
- 4 See David Rodin, *War and Self-Defense* (Oxford: Oxford University Press, 2002), pp. 88–90; Michael Otsuka, 'Killing the innocent in self-defense', *Philosophy & Public Affairs* 23,1 (1994): 74–94; and McMahan 2009 op. cit., pp. 167–73.
- 5 Jonathan Quong, 'Killing in self-defense', *Ethics* 119,3 (2009): 507–37; and Helen Frowe, 'Equating innocent threats and bystanders', *Journal of Applied Philosophy* 25,4 (2008a): 277–90.
- 6 See Jeff McMahan, 'The basis of moral liability to defensive killing', *Philosophical Issues* 15 (2005): 386–405; Kai Draper, 'Defense', *Philosophical Studies* 145, 1 (2009): 69–88; and Rodin op. cit., pp. 77–9. Those who accept such views may permit one to kill a nonresponsible threat in order to save many lives. As I indicate below, however, I will set aside these cases.
- 7 Helen Frowe offers a hybrid view on which culpability is necessary for one to be liable to defensive force, though an innocent person is permitted to kill a non-liable threat who will otherwise kill him. One

- implication of her view is that a non-labile threat is permitted to employ counter-defence to preserve his own life. See Helen Frowe, 'A practical account of self-defence', *Law and Philosophy* 29,3 (2010): 245–72. Quong (op. cit., pp. 520–1) agrees that, for instance, Falling Person is permitted to shoot you before you shoot him.
- 8 Whitley Kaufman complains that critics of the traditional view, including Rodin and Otsuka, end up '*in practice* with a position virtually indistinguishable from the traditional view'. See Whitley Kaufman, 'Self-defence, innocent aggressors, and the duty of martyrdom', *Pacific Philosophical Quarterly* 91,1 (2010): 78–96, p. 94, emphasis added. It is not clear why we should regard this as a problem for critics of the traditional view, however. To consider an analogy, it is at least possible that, *in practice*, the most plausible form of consequentialism will end up being very close to the most plausible form of deontology. But this would not show that there is no important difference between consequentialism and deontology, nor would it provide an objection to either theory.
  - 9 See Nancy Ann Davis, 'Abortion and self-defence', *Philosophy and Public Affairs* 13,3 (1984): 175–207, pp. 194–6; cf. Quong op. cit., pp. 516–8.
  - 10 Otsuka op. cit., pp. 84–5.
  - 11 Warren Quinn, 'Actions, intentions, and consequences: The doctrine of double effect', *Philosophy & Public Affairs* 18,4 (1989a): 334–51, p. 344; and Quong op. cit., p. 525.
  - 12 Quinn 1989a op. cit., p. 344.
  - 13 Frowe 2008a op. cit., p. 281.
  - 14 See Judith Jarvis Thomson, 'Self-defence', *Philosophy and Public Affairs* 20,4 (1991): 283–310, p. 291; and Quong op. cit., p. 526.
  - 15 This problem is also noted in connection with a different example in McMahan 2009 op. cit., pp. 171–3.
  - 16 Quong op. cit., p. 528.
  - 17 Quong op. cit., p. 530, emphasis added.
  - 18 Quong op. cit., p. 529, n. 48.
  - 19 Quong's counterfactual test may face problems in cases involving overdetermination. Suppose, for example, that two people begin shooting at you. One gunman is culpable, and the other is nonresponsible. (Imagine that your enemy placed a gun in the nonresponsible gunman's hand just as he began to have seizure-induced muscle spasms in his trigger finger.) The nonresponsible gunman is standing out in the open, where you can easily shoot him, while the culpable gunman is hiding behind a car, which protects him from your defensive action. The only way for you to save your own life is to shoot the nonresponsible gunman in such a way that his body will block the culpable gunman's line of fire. In this case, you do profit from the presence of the nonresponsible gunman: if he were not present, you would be unable to save yourself from the culpable gunman. But it seems clear that the traditional view should still permit you to kill the nonresponsible gunman in this way. I am not sure whether Quong's counterfactual test could be modified to accommodate this point, but I will not pursue this issue further, especially since I am focusing on cases involving just one threat and one prospective victim. A somewhat similar example is offered (for different purposes) in McMahan 2005 op. cit., p. 391.
  - 20 For a similar example, see Philippa Foot, 'The problem of abortion and the doctrine of double effect' in B. Steinbock & A. Norcross (eds) *Killing and Letting Die*, 2nd edn. (New York: Fordham University Press, 1994), p. 276.
  - 21 Although Quong does not consider this sort of example, such a response may be suggested in Quong op. cit., pp. 531–2.
  - 22 Quong (op. cit., p. 532, n. 52) claims that it would be permissible for you to kill another person by wobbling her off a bridge if the bridge is your own property and you can save your life only by wobbling the bridge.
  - 23 Following Judith Jarvis Thomson, I consider one to have infringed a right just in case one did not accord another something that he had a right to. I consider one to have violated a right just in case one *wrongly* infringed a right. See Thomson, *The Realm of Rights* (Cambridge, MA: Harvard University Press, 1990), p. 122.
  - 24 I am indebted to Kai Draper for pointing out that this sort of case poses problems for Quong's position.
  - 25 Frowe 2008a op. cit., p. 285.
  - 26 Helen Frowe, 'Threats, bystanders and obstructors', *Proceedings of the Aristotelian Society* 108,3 (2008b): 365–72.
  - 27 Frowe 2008a op. cit., p. 285; see also Frowe 2010 op. cit., pp. 270–1. Frowe (2008a op. cit., p. 285) also counts a person as a direct threat if he 'is attached to an object that will kill' the victim.
  - 28 Frowe 2008a op. cit., p. 285.
  - 29 Thomson 1991 op. cit., pp. 288–9.

- 30 Thomson 1991 op. cit., p. 292.
- 31 Suzanne Uniacke, *Permissible Killing: The Self-Defence Justification of Homicide* (Cambridge: Cambridge University Press, 1994), p. 178, emphasis added.
- 32 Thomson 1991 op. cit., pp. 288–9.
- 33 This does not necessarily mean that an agent can never do harm by staying still. I might stand in your path, completely stationary with my foot extended, in order to trip you. In this case, I do harm. But it seems that I have done harm not merely by staying still, but instead by deliberately or knowingly placing myself in your path. Nothing similar is true of Stationary Person in *Fall (Inverse)*. Indeed, for the purposes of the example, we need not even suppose that Stationary Person chose to be in his present location; perhaps he was placed there by some evil-doer. I am indebted to an anonymous referee for this journal, who pointed out the need for greater clarity on this point.
- 34 See Warren Quinn, ‘Actions, intentions, and consequences: The doctrine of doing and allowing’, *The Philosophical Review* 98,3 (1989b): 287–312, p. 287; and Otsuka op. cit., pp. 78–9.
- 35 See the discussion of the ‘causer pays’ principle in Cheyney Ryan, ‘Self-defense, pacifism, and the possibility of killing’, *Ethics* 93,3 (1983): 508–24, pp. 516–7, though Ryan (p. 519) rejects this principle.
- 36 Quong op cit., p. 508. This definition is similar to that offered by Otsuka (op. cit., p. 75).
- 37 Of course, neither Speedy nor Stationary Person will be *the* cause of your death. In the former case, whatever started the train in motion will also be a cause of your death, and in the latter case the evil-doer who pushed you over the cliff will also be a cause of your death.
- 38 See, for instance, Kadri Vihvelin & Terrance Tomkow, ‘The dif’, *Journal of Philosophy* 102,4 (2005): 183–205, esp. p. 198. Of course, even if Speedy will be a cause of your death, this does not necessarily mean that he has acted wrongly.
- 39 Philippa Foot argues that to do harm is to initiate a harmful causal sequence or to sustain such a sequence when it otherwise would have stopped. See Philippa Foot, ‘Killing and letting die’, in B. Steinbock & A. Norcross (eds) *Killing and Letting Die*, 2nd edn. (New York: Fordham University Press, 1994), pp. 282–3. See also Ryan op. cit., p. 514.
- 40 See Rodin op. cit., p. 82.