

## The Moral Responsibility Account of Liability to Defensive Killing<sup>1</sup>

*Michael Otsuka*

Sometimes human beings pose threats to the lives of other human beings. They may or may not be blameworthy for doing so. Among those who are blameless, some may be blameless because they are not morally responsible for being a threat, whereas others may be blameless in spite of their responsibility. On what has come to be known as the '**moral responsibility account**' of liability to defensive killing, it is such responsibility, rather than blameworthiness, for threatening another that renders one liable to defensive killing. Moreover, one's lack of moral responsibility for being a threat immunizes one from liability to defensive killing. About twenty years ago, I offered an early formulation and defense of such an account.<sup>2</sup> I would like to renew my defense of the moral responsibility account here.

It will help to provide an example of a blameless but morally responsible threat. I offered the following **example of a foreign dignitary** who is blameless but morally responsible for being a threat to your life:

Imagine that you extend your hand to shake the hand of some foreign dignitary at a reception. Unbeknown to you, a third party projects a stunningly realistic holographic image of a pistol onto your hand. The dignitary, who is accustomed to threats on her life, sees the hologram, forms the justified belief that you are about to assassinate her, and coolly draws a pistol in order to shoot you down in self-defense. (91)<sup>3</sup>

---

<sup>1</sup> Thanks to Jonathan Quong for comments.

<sup>2</sup> In Michael Otsuka, "Killing the Innocent in Self-Defense," *Philosophy and Public Affairs* 23 (1994): 74-94. I didn't, however, call this the 'moral responsibility account.' The name is, I think, due to Jeff McMahan.

<sup>3</sup> All such parenthetical page references are to "Killing the Innocent in Self-Defense."

The dignitary is morally responsible because she is of sound mind, in control of her actions, and aware of their likely consequences. But she is not blameworthy for trying to kill you, since she acts from the justified (albeit false) belief that she is fending off an assassination attempt.

Contrast this with the following example of someone who is neither blameworthy nor morally responsible for posing a threat to your life: a person who has been transformed by an unpredictable tornado into an unconscious **human projectile** that is now lethally falling towards you. You cannot step safely to one side because “philosophers have arranged” that the same tornado has gently deposited you at the bottom of a well.<sup>4</sup>

Whereas I argued that you would be justified in killing the foreign dignitary in self-defense, I also argued that it would be wrong to use lethal force to defend yourself against the human projectile. I argued that it would be wrong to kill the latter on grounds that such a threat could not be morally distinguished from an innocent bystander. I also argued that it is wrong to kill an innocent bystander in self-defense, at least when the killing involves the initiation rather than the redirection of a lethal chain of events.

In Section I of this paper, I renew my defense of the claim that it is impermissible to kill a nonresponsible threat such as the wind-blown human projective in self-defense. In Section II, I renew my defense of the claim that it is permissible to kill a blameless but morally responsible individual such as the foreign dignitary in self-defense.

## **I. The case against killing a nonresponsible threat**

---

<sup>4</sup> The phrase in quotation marks is Philippa Foot's. See Foot, "The Problem of Abortion and the Doctrine of the Double Effect." Robert Nozick arranged the venue of the bottom of the well. See Nozick, *Anarchy, State, and Utopia*, pp. 34-5. And various other philosophers added the wind.

As I have mentioned above, I argued that it is impermissible to kill a nonresponsible threat such as the human projectile on grounds of his moral equivalence to an innocent bystander. My main argument for this equivalence invoked the following trio of cases:

**First imagine that an innocent person** is lying alongside the path of a runaway trolley car. Unless you hurl at that trolley a bomb [i.e., a powerful grenade] that you know will also kill the innocent person, the trolley will run you over. (85)

I argued that it would not be permissible to foreseeably kill such an innocent bystander in self-defense in this case. I then asked the reader to imagine:

**a second case in which the same person is trapped** inside a runaway trolley car. Unless you hurl a [powerful grenade] that will destroy the trolley, and hence also the innocent person, the trolley will run you over before coming to a gentle stop. (85)

I maintained that "If doing that which foreseeably will kill the person is impermissible in the case in which the person is alongside the trolley, then I do not see how it could be permissible if the person is inside the trolley. Changing the location of the person should not make any moral difference" (85).

I then maintained that "If, however, it is impermissible to destroy the trolley inside of which a person is trapped, then I do not see how it could be permissible in **a third case to vaporize a falling body that itself constitutes a [nonresponsible] Threat**" (85).

In *The Ethics of Killing*,<sup>5</sup> Jeff McMahan noted that some would resist my argument by maintaining that, in Case I, it is not in fact impermissible to hurl the grenade at the trolley in self-defense, even though one foresees that this will kill the innocent bystander alongside the trolley.<sup>6</sup> I had argued that the claim that such a foreseeable killing is impermissible gains support by its coherence with claims regarding the impermissibility of foreseeable killings in other cases, such as **one inspired by Philippa Foot that I invoked in my article** in which “you must drive over one recumbent Bystander in order to prevent yourself from being swept away by a rising tide” (77). Jeff, however, resisted my assimilation of the grenade-hurling case to the Foot-inspired case in which you must run someone over to survive. He proposed that, even though the killings are analogous insofar as they are both foreseen rather than intended, the grenade-hurling killing is less morally problematic than the killing in the driving-over case for the following reason: the harm to the person is, in Jonathan Bennett’s terminology, *causally upstream* the good of saving your own life in the driving-over case. In other words, the harm to the person is on (or forks off of) the causal stream that flows to the good of the saving of your life. But the harm to the person is not causally upstream in this sense in the grenade-hurling case. Rather it is, so to speak, *causally sidestream* the good insofar as it is a side-effect of the very event – the grenade’s obliteration of the trolley – that constitutes the good.<sup>7</sup> Causally upstream killings appear, intuitively, to be morally more problematic than otherwise similar killings that are instead sidestream or downstream the good.

**As a friendly amendment, Jeff then proposed the replacement of my Case I** with a different case in which the killing of an innocent

---

<sup>5</sup> OUP, 2002.

<sup>6</sup> He discusses my argument on pp. 407-9.

<sup>7</sup> Jeff suggests that the grenade killing is causally *downstream* the good. But I think it’s better characterized as ‘sidestream’.

bystander is more clearly impermissible. On Jeff's amendment, your killing of an innocent bystander is a means of triggering a bomb to go off that will destroy the trolley that is threatening you.

A problem with Jeff's friendly amendment is that here the innocent bystander is someone whose presence you opportunistically exploit in order to save your life, and you would not have been able to save your life if this innocent person had been absent from the scene. That feature renders this case disanalogous from the case in which you vaporize a falling nonresponsible threat since, if that person had been absent from the scene, you would have been completely safe from harm.<sup>8</sup> Killing someone as a means of detonating a bomb is, in Warren Quinn's helpful terminology, a case of "opportunistic agency," whereas the killing of the falling projectile is a case of "eliminative agency." Since opportunistic agency is, other things equal, more morally problematic than eliminative agency, one cannot straightforwardly appeal to the wrongness of killing an innocent bystander as a means of detonating a bomb in Jeff's revised Case I in order to establish the wrongness of killing a falling threat. So one cannot move from the impermissibility of killing in Jeff's revised Case I to the impermissibility of killing in my Case III involving the falling threat.<sup>9</sup>

**Jeff's revision of my Case I therefore amounts to overkill**, so to speak. If the killing of an innocent bystander in my original Case I is not sufficiently clearly morally problematic for my purposes, Jeff revises that case into a case that is, so to speak, *too* morally problematic for my purposes. **I need to find a 'Goldilocks' case in between the two**: one that contains just the right balance of moral features. By this I mean a case where killing in self-defense is intuitively impermissible by virtue of features

---

<sup>8</sup> Helen Frowe identifies this problem with Jeff's example in "Equating Innocent Threats and Bystanders," *Journal of Applied Philosophy*, 25 (2008): 277-90 at p. 283.

<sup>9</sup> See Jeff's discussion on p. 170 of *Killing in War* (OUP, 2009), for similar remarks regarding the relevance of Quinn's eliminative/opportunistic distinction to the topic of killing in self-defense.

that render it analogous to the killing of a nonresponsible threat rather than by virtue of features that render it morally more problematic than the killing of such a threat.

In my search for the perfect Goldilocks case, I shall first consider a case involving an **innocent obstructor** that Jeff has more recently settled on, in *Killing in War*, in order to defend the claim that it is impermissible to kill a nonresponsible threat. Unlike a threat, an innocent obstructor poses no danger of killing you. But, like a threat, and unlike a bystander, his presence poses a problem for you: namely, this person's live body is in the way of what you need to save your life. In Jeff's case, you must cross "**a high, narrow, and wobbly public bridge**" in order to escape a murderer.<sup>10</sup> Unfortunately, the only way for you to get across this bridge is to clear a path for yourself by **toppling someone else off the bridge** who is innocently minding his own business. Intuitively, such a killing is impermissible. Moreover, **Jeff maintains that "there is no difference in the mode of agency involved** in killing [such an innocent obstructor] and that involved in killing a Nonresponsible Threat." He also maintains that there are no other grounds by which we can morally distinguish the killing of the one from the killing of the other.<sup>11</sup> So he infers the impermissibility of killing a nonresponsible threat from the impermissibility of killing such an innocent obstructor.

There is, however, a problem with this line of argument. The problem, in fact, **parallels the one that Jeff identified when he criticized the analogy I drew between my grenade-hurling Case I and the Foot-inspired case in which you must drive over someone in order to save your life.** Recall that Jeff noted that the driving over the one is *causally upstream* the good, whereas the killing of the one with the shrapnel from the grenade that also eliminates the trolley that is barreling towards you is not

---

<sup>10</sup> *Killing in War*, p. 171.

<sup>11</sup> See *Killing in War*, pp. 171-3.

causally upstream the good. Recall also that causally upstream killings appear, intuitively, to be morally more problematic than otherwise similar killings that are not causally upstream. But now note that the driving-over case involves an innocent obstructor: the body of the recumbent person on the road is in the way of what you need to save your life. Hence in both this case and in Jeff's bridge case, there is an obstructor who is blocking your path to the good, and you have to eliminate these obstructions in order to gain access to the good in question.

**The problem is that the apparently morally significant distinction between killings that are causally upstream the good, and those that aren't, can be deployed in order to distinguish the killing of the innocent obstructor on the bridge from the killing of a nonresponsible threat.** The killing of this obstructor is causally upstream the good of the elimination of a threat to your life. By contrast, the vaporization of a falling threat is not causally upstream the achievement of the good of the elimination of the threat to your life. Rather, it is one and the same event as this elimination. So Jeff's objection to my appeal to the driving-over case to defend the impermissibility of the hurling of a grenade in Case I turns out also to be an objection to his own appeal to the obstructor on the bridge in order to defend the impermissibility of the killing of a nonresponsible threat. The objection is that it is not clear that one is entitled to appeal to the impermissibility of an upstream killing in order to establish the impermissibility of an otherwise similar but non-upstream killing. This is because these killings differ in the apparently morally significant respect that one is upstream and the other is not.<sup>12</sup>

---

<sup>12</sup> In *The Ethics of Killing*, Jeff maintains that "there seems intuitively to be a moral difference between the mode of agency in Foot's [driving over] case and that in [Otsuka's grenade-hurling Case I]." (*Ethics of Killing*, p. 408.) But we can now see that this claim is at odds with his claim, in *Killing in War*, which I quoted earlier, that "there is no difference in the mode of agency involved in killing [the innocent obstructor on the bridge] and that involved in killing a Nonresponsible Threat."

We need to turn to a different sort of case in order to establish the impermissibility of the killing of a nonresponsible threat.<sup>13</sup> We need a case of an obstructor whose elimination is *not* causally upstream the securing of the good of one's life. Rather, it is a case in which his elimination, or rather his displacement, is *itself* the securing of the good of the saving of one's own life. There is a case of Judith Thomson's that seems to fit the bill perfectly. Perhaps it is the fabled Goldilocks case:

**Alcove Case:** "A subway is headed toward you. ... There is a small alcove in the wall near you, but there is another [individual] already in it. You can force your way into the alcove, thereby crushing him to death."<sup>14</sup>

**It is plausible to maintain that, in this case,** the crushing of the one is not causally upstream the securing of the good of refuge from the subway. Rather, it is one and the same event as the securing of this good. It is, in this respect, akin to the vaporization of the falling threat insofar as his vaporization just *is* the good of the elimination of the threat, rather than being causally upstream this good. I grant that there is the following respect in which the crushing of the one could be described as causally upstream the good: it is something that you need to accomplish in order to free up life-saving space to take refuge in the alcove. I can grant this because the vaporization of the falling threat could, in the same respect, be described as causally upstream the good, since that too is something that you need to accomplish in order to save your life. Whether one categorizes both killings or neither as causally upstream the good, it remains the case that the

---

<sup>13</sup> In his earlier writings on self-defense, Jeff offered a different case-based argument, involving trapped miners, in order to try to establish the impermissibility of the killing of a nonresponsible threat. But, as I show in the appendix to this paper, this case also suffers from an upstreamness objection.

<sup>14</sup> Thomson, "Self-Defense," p. 291.



**crushing of the person in the alcove is morally on a par with the vaporization of the falling threat:** there are no morally relevant differences between the two cases.<sup>15</sup> Moreover, the crushing of the person in the alcove is impermissible. So, therefore, is the vaporization of the falling threat.<sup>16</sup>

**Jonathan Quong would resist the above argument for the impermissibility of killing a nonresponsible threat.** He would try to morally distinguish the alcove and the falling threat cases in the following way. First, Jon would maintain that crushing the one in the alcove is impermissible for much the same reason that it is impermissible to grab an innocent bystander to make use of him as a shield from an oncoming threat. In each case, the person's body or the physical space he occupies is useful to you. Hence, you "take advantage of, or exploit, the body or the physical space of someone else." Moreover, in each case, the person has "a rightful claim" -- i.e., a "claim-right" -- to this body or space.<sup>17</sup> When you vaporize a falling threat, by contrast, you do not take advantage of or exploit his body or his physical space, as you make use of neither.

I agree with Jon that whether the bystander or the obstructor has a rightful claim to the body or space that he possesses is relevant to the

---

<sup>15</sup> Here I am asserting, rather than arguing for, such morality parity. But in "Killing the Innocent in Self-Defense," I rejected a number of arguments that could be offered on behalf of the claim that the killing of such a falling threat is morally distinguishable from the killing of the person in the alcove.

<sup>16</sup> If one reason why it is intuitive that crushing the one in the alcove is impermissible is that this is an 'up close and personal' killing, note that people would condemn as impermissible a comparably up close and personal killing of a falling threats: e.g., you need to reach your hands up and deflect the falling threat over a cliff to prevent him from landing on your head and killing you.

<sup>17</sup> Quong, "Killing in Self-Defense," *Ethics*, pp. 528-9. Quong explicitly addresses a different alcove case, also due to Thomson, in which you need to grab the person in the alcove and throw him onto the track in order to gain lifesaving access to this alcove. Insofar as my aim of demonstrating the impermissibility of killing an innocent threat is concerned this, alcove case suffers the same problem as Jeff's bridge and miner cases: the harming of the one is causally upstream the good in a manner that contrasts with the vaporization of a threat. This case is also less good for Quong's purposes for a reason which I indicate in a footnote below.

question of the moral permissibility of killing this person in self-defense. To see the relevance, take first the case of the bystander that you can use as a human shield. A javelin is hurtling towards you. You reach over to grab this person's body and use it as a shield. The person could legitimately complain as follows: **"I'm sorry, but this is my body and not yours. So keep your hands off!"**

Now consider a contrasting case of Jon's in which the body or space in question is rightfully yours rather than his:

**Meteor:** A small meteor is falling toward you and will kill you if it lands on you. The only safe place where you can avoid the meteor is your very tiny one-person car. But there is already someone in your car—this person was placed there without their consent by some third party. You could, however, pull them out of the car, thereby ensuring they will die, so you can get inside to safety.<sup>18</sup>

Jon maintains that it would be permissible to pull this person from the car and occupy it yourself, on grounds that "it is, after all, your car" and not his. He deems this permissible even if this would clearly involve a killing rather than a withdrawal of aid because "the only way to remove them from the car is to throw them out the door and off the edge of the cliff."<sup>19</sup> I'm not sure whether I share Jon's intuition that this is permissible. I am, however, more willing to go along with an analogous claim that it would be permissible for you to jump into the car, thereby crushing the occupant, if that were the only way for you to gain refuge from the meteor.<sup>20</sup> It is also worth adding that, if this were *his* car, rather than yours, then it would clearly be

---

<sup>18</sup> P. 527.

<sup>19</sup> Ibid.

<sup>20</sup> So we can see now that a further reason why Quong would have been better served by the crushing rather than the removing version of Thomson's alcove case is that it is more intuitive, in the analogous car case, that it is permissible to kill the person in the car.

impermissible for you to take refuge in his car, either by removing him from the car or by crushing him to death.

How does this bear on Thomson's alcove case? If the alcove is his rather than yours because he rightfully owns it, then it is, by parity of reasoning, impermissible for you to take refuge there, thereby crushing him. But what if neither you nor he has a legitimate title deed to the alcove? What if we stipulate that the alcove is neither your nor his private property but rather is public property? We are, after all, speaking of an alcove on a metropolitan subway platform in Thomson's example. Here it would seem that the person who happens already to be in the alcove has no right over the space to ground the claim that it is impermissible for you to occupy his space, thereby crushing him to death.

Jon thinks otherwise. He maintains that the person already in the alcove has a claim-right to the space along the lines of the following individual in a public park:

Suppose Albert and Betty are in a public park, and **Albert desires to get a better view of the lake but can only do so if Betty moves from her present location** where she is enjoying a picnic. Albert cannot simply move Betty, even if this causes her no harm, and even if he is somehow able to do this without touching Betty (thus avoiding violating any claims she might have against nonconsensual touching). Betty has a presumptive claim-right to her location even though it is a public park, and even though it is more or less arbitrary that Betty arrived at that particular spot first.<sup>21</sup>

Here is one problem with such an appeal to the existence of rightful claims over the world as the explanation for the impermissibility of killing in

---

<sup>21</sup> P. 528.

the alcove case. **Suppose that it is a two-year-old child** rather than a full-grown adult whom you would crush if you jump into the alcove to escape the oncoming trolley. Intuitively, it is at least as clear that it is impermissible to crush such a child as it is to crush a full-grown adult. On Jon's account of permissibility, this would be by virtue of the fact that this child has acquired a rightful claim to the space she occupies. One might, however, doubt that such young children can have rightful claims to the space they occupy, especially given that Jon's rationale for the existence of such claim-rights is "grounded in a view of persons as self-directing agents."<sup>22</sup> But even if we restrict our attention to rational full-grown adults, Jon's account can be challenged, as I shall now show, on grounds that people in public spaces such as parks and alcoves lack the claim-right to which he appeals.

Jon notes that "**Betty's rights [to her spot in the park]** may sometimes be permissibly overridden if the benefits to some other person or persons are substantially greater than any interest of Betty's in remaining where she is." But he maintains that "we cannot simply ignore Betty's rights whenever there might be some small overall good in doing so."<sup>23</sup> Suppose, however, that Betty occupies the only spot on in the park with a view of the lake. She sits there hogging the space from sunup to sundown. Would it not be appropriate at some point for her to share this spot with others? If asked to do so, wouldn't it be unconvincing for her to reply: "No, I'm sorry, but I have a rightful claim to this spot because I got here first." And **if she happened to be on a Lazy Susan**, would it not eventually be permissible for Albert to rotate it so that he can have a look at the view, even though he thereby moves Betty from her spot?

Now this case differs from the alcove case insofar as it is possible to share the enjoyment of the view equally between the two, whereas at most one person will be able to enjoy refuge in the alcove from the oncoming

---

<sup>22</sup> Ibid., p. 529.

<sup>23</sup> P. 528.

subway train. Suppose, however, that as you go to seek refuge in the alcove from the train that you know will be arriving in five minutes, you discover that someone else is already there. What if you propose that the two of you determine by fair lot whether he gets to remain in the alcove, or whether you get to seek refuge there? Suppose that he replies as follows: "I'm sorry, but I got here first. So it's my alcove, to which I have a rightful claim, and to which you have none." Here, he would speak falsely. Even though you got there second, you have a legitimate claim to an equal chance of refuge in the alcove.

Consider the following analogy: A cruise ship in the Caribbean has sunk, and you and someone else are treading water. Then, as luck would have it, a life preserver ring floats by. It is clear from the markings that it is not from this particular cruise ship. It's nearer to the other person, so **he simply grabs it and places it around his neck. "Wait a minute!" you protest. "Shouldn't we flip a coin to determine who gets this?"** Suppose that he replies: "I'm sorry. But I have a rightful claim to this preserver, because I grabbed it first." I think the right response would be: "No. I'm sorry, but I have as much a claim to this as you do." Suppose that you then toss a fair coin, declaring while it's in the air that heads you get the preserver, and tails he does. If it then lands heads, **I think you are entitled to wrest the life preserver from him**, especially if you can do so without even laying a hand on him.

I'm less inclined to affirm the permissibility of your jumping into the alcove and crushing the first occupant after winning a fair coin toss that you unilaterally impose. But this is not because I think the first occupant of the alcove has any greater claim to it than the first occupier of the life preserver has over it. I think claims over each of the alcove and the preserver extend to nothing stronger than an equal shot at it. The reason why I am less inclined to affirm the permissibility of your jumping into the alcove and

crushing the person is that this would constitute a violation of the person's rights of self-ownership.

Consider a contrasting case in which you could rotate someone else out of, and yourself into, the alcove by turning a Lazy Susan on which he and you are standing. I believe that you would be entitled to do so after a winning coin toss that you unilaterally impose, as here you are able to remove him from the alcove without infringing his self-ownership. I grant that not every case involving the rotating of a Lazy Susan is a case in which no rights of self-ownership are infringed. If, for example, you were to impale someone on a spear by moving him into it by rotating the Lazy Susan on which he stands, it is hard to resist the claim that you would thereby infringe his self-ownership. But we should contrast this with a case in which turning your Lazy Susan places someone else on the outside rather than the inside of your house. **Imagine that the Lazy Susan is akin to the revolving doors at the entrances to buildings that are installed to prevent cold air from rushing in.** If you thereby make it the case that a stranger who has taken refuge inside your house is now **outdoors in the cold where he will freeze to death**, rather than indoors, you don't infringe his self-ownership. Rather, you merely withdraw life-saving aid. The same goes if, by turning your Lazy Susan, you move the stranger outdoors, **where he will be attacked by the wild animals** that roam the streets outside your house. And I think the same goes if you turn the Lazy Susan, thereby **rotating him out of the alcove and into the path of the oncoming train.**

The upshot of this discussion is that it is a right of self-ownership not to be imposed upon, rather than a right not to have made use of something to which one has a rightful claim -- whether it be one's body or the space one occupies -- which explains the impermissibility of jumping into the alcove and crushing the person who has taken refuge there. The right not to

be imposed upon here is a right against being the object of harmful eliminative agency. This right obtains even when the harmful eliminative agency is not causally upstream the good, insofar as the crushing of the one and the occupation of the refuge are one and the same. The morally impermissible crushing of the person in the alcove is morally on a par with the vaporization of the falling threat. Hence the latter is impermissible.

## **II. The case for killing a blameless but morally responsible threat**

I would now like to turn to a discussion of cases in which it is permissible rather than impermissible, on the moral responsibility account, to kill another in self-defense. In particular, I would like to consider a class of cases in which it is permissible to kill somebody because she is morally responsible even though she is also blameless. **Recall, for example, the case involving a foreign dignitary and a hologram**, with which I opened this paper. If the dignitary shoots in self-defense, she would be doing something objectively wrong: she would be killing a nonresponsible person who is not liable to be killed, and whose killing cannot be justified on grounds apart from liability. But such a killing would be subjectively morally justifiable, since she would be acting on the justified (but false) belief that she is defending herself against an attack by a culpable assassin.

Is the dignitary liable to be defensively killed, if this is the only way to save the life of the person who has innocently reached out to shake her hand? May one kill such a morally responsible but blameless active threat in order to save the life of the innocent whom she threatens?

As far as I am aware, **Kai Draper was the first person to pose such a question about morally responsible but non-culpable active threats**. In an article called "Fairness and Self-Defense," which was

published in 1993, Draper asks us to imagine a case along the lines of the dignitary, in which:

I live in a dangerous area and so I justifiably buy a hand gun for protection. But this leads to tragedy when I begin shooting at my neighbor because I reasonably but mistakenly think that he poses a threat to my life. He can save himself only by shooting me. (p. 84)

Draper argues that it would be justifiable for my neighbor to shoot me in self-defense. His argument is very brief but to the point:

In this case I know when I buy the gun that guns are dangerous and that I could potentially pose a threat to an innocent person. But I justifiably choose to create this risk for the sake of my own security. Thus, when this risk is realized in my attack on my neighbor, it seems fair that I should sustain the costs.<sup>24</sup> (ibid.)

A year later, and in ignorance of Draper's account, **I offered a justification for killing such a morally responsible but non-culpable active threat** that also appealed to considerations of fairness:

When one is in possession of rational control over such a dangerous activity as the shooting of a gun at somebody, it is not unfair that if the person one endangers happens to be innocent, one is by virtue of engaging in such dangerous activity stripped of one's moral immunity from being killed. *A responsible agent takes a gamble by placing this moral immunity on the line when she engages in such avoidable risky activity.* (91, emphasis added)

---

<sup>24</sup> Ibid.



As this passage indicates, my own explanation for the permissibility of defensively killing such morally responsible but blameless active threats as the dignitary was as follows: the charge that it would be unfair if people such as the dignitary were liable to be killed is answered by pointing to the fact that they gambled with their moral liability to being defensively killed by engaging in such activity that they knew to be risky. **I was here drawing an implicit analogy between the bad moral luck of becoming morally liable to be killed and Ronald Dworkin's notion, within the realm of distributive justice, of bad option luck.** Option luck, as Dworkin defined it, is "a matter of . . . whether someone gains or loses through accepting an isolated risk he or she should have anticipated and might have declined." Brute luck, by contrast, is "a matter of how risks fall out that are not in that sense . . . gambles." Dworkin illustrated this distinction as follows: "If I buy a stock on the exchange that rises, then my option luck is good. If I am hit by a falling meteorite whose course could not have been predicted, then my bad luck is brute (even though I could have moved just before it struck if I had any reason to know where it would strike)."<sup>25</sup> In an article entitled "Moral Luck: Optional, not Brute," I made explicit this analogy between the bad moral luck of gambles with one's moral liability and the bad option luck of financial gambles.<sup>26</sup> By drawing on this analogy, **one can answer a charge of unfairness when someone voluntarily chooses to gamble with his moral worth**, so to speak, by knowingly engaging in risky activity that might imperil the innocent. The person who tries to shoot someone who turns out to be innocent is morally unlucky in much the same way that someone who gambles on the stock market and loses is financially unlucky. The upshots of such gambles are not unfair because they are a matter of bad option luck rather than brute bad luck.

---

<sup>25</sup> Dworkin 1981: 293.

<sup>26</sup> *Philosophical Perspectives* (2009).

Let us now apply this analysis to the following case of Jeff's {called the ***Cell Phone Operator*** but which, in my PPT slide, I have renamed 'Press "send" for "SHTF"'}:

***The Cell Phone Operator:*** A man's cell phone has, without his knowledge, been reprogrammed so that when he next presses the "send" button, the phone will send a signal that will detonate a bomb that will then kill an innocent person.<sup>27</sup>

The cell phone operator is a responsible agent. He is, one might say, as responsible as any other normal adult human being for the choice of buttons he pushes on his phone. But Jeff plausibly maintains that the cell phone operator is not responsible for posing a threat to another by pressing "send." He is not responsible for posing a threat because he has no way of knowing that doing so will kill someone. In Jeff's words "he is nonculpably and invincibly ignorant *that he poses any kind of threat to risk of harm to anyone.*"<sup>28</sup>

If, moreover, we apply the option-luck/brute-luck distinction to this case, it is clearly an instance of brute luck rather than option luck, since, in pressing "send," he has not thereby accepted "an isolated risk he or she should have anticipated and might have declined." Of course it is as the result of his choice to press "send," which he could have refrained from making, that he ends up killing the one. But such avoidability does not distinguish this case from Dworkin's paradigmatic meteorite example of brute luck in which we can imagine that it is as the result of a choice to sit at this table rather than the next one that I end up being hit by the meteorite.

The meteorite example also serves to answer one line of objection to the distinction that Jeff and I wish to draw between cases such as the

---

<sup>27</sup> *Killing in War*, p. 165.

<sup>28</sup> *Killing in War*, p. 168.

dignitary on the one hand and the cell phone operator on the other hand. One might argue that *anything* we do involves a gamble, because we never know what weird and wild risks might arise from our choices. We have to be ever vigilant to the possibility that whatever we do might harm others by however improbable a causal chain. But the fact that every action is, in some sense, a gamble insofar as it exposes one to risk of some sort, including a gamble to sit at a table that ends up being crushed by a meteorite – that fact doesn't obliterate the moral significance of the distinction between a loss that is the upshot of one's gambling everything on a roulette wheel and a loss that is the upshot of one's having chosen a table in the path of a meteorite. One couldn't credibly demand compensation for one's losses at roulette on grounds that one has just as strong a claim of compensation for bad luck as the person who was hit by the meteorite. Similarly, the fact that there's some chance that pressing "send" will kill an innocent doesn't obliterate the distinction between pressing "send" and firing a bullet at someone.

I would like to close by considering another case of Jeff's. Like his cell phone case, this one involves an unintended harm, whereas my case of the foreign dignitary involves an intention to harmfully incapacitate another. But unlike the cell phone case, in this case one is aware of the fact that one's activity poses a real danger to others:

***The Conscientious Driver:*** A person who always keeps her car well maintained and always drives carefully and alertly decides to drive to the cinema. On the way, a freak event that she could not have anticipated occurs that causes her car to veer out of control in the direction of a pedestrian.<sup>29</sup>

---

<sup>29</sup> *Killing in War*, p. 165

Jeff offers the following commentary on this case, with which I am in agreement:

Although she does not intend to harm anyone, she does know that her action carries a small risk of causing great though unintended harm. Although her act is of a type that is generally objectively permissible, and although she has taken due care to avoid harming anyone, she has had bad luck: the risk she knew her act carried has now, improbably and through no fault of her own, been realized. Because she knew of the small risk to others that her driving would impose, and because she nonetheless voluntarily chose to drive when there was no moral reason for her to do so[,] ... she is morally liable to defensive action to prevent her from killing an innocent bystander.<sup>30</sup>

**In an article entitled “Liability to Defensive Harm,” Jonathan Quong** concurs with Jeff and me that the dignitary has, by taking aim at an innocent, rendered himself liable to being defensively killed. Yet he rejects the view that the conscientious driver has also rendered himself liable. He rejects our moral responsibility account of liability in favor of a different one, which tracks an interesting feature that differentiates the dignitary from the conscientious driver. In raising her pistol to shoot the person whom she believes to be threatening her life, the dignitary acts as if the person whom she confronts lacks rights that he in fact has. And Jon maintains that:

To treat others as if they lack moral rights against having harm imposed is a grave matter, and so it is plausible to suppose that when we act in this way, we must accept a certain substantive responsibility for our actions. If we treat others as if they are liable to harm, it

---

<sup>30</sup> Ibid., p. 166.

seems only fair to suppose that we may become liable to defensive harm should that judgment be mistaken.<sup>31</sup>

By contrast, the conscientious driver never acts on the assumption that anyone lacks rights that he in fact has. The justification of our traffic laws to which he conforms, which appeals to a calculation of the costs and benefits of driving, including the cost of accidental deaths, does not depend on any assumption that people lack rights not to be harmed that they in fact possess. Nor, of course, does any individual driver hurtle accidentally out of control towards a pedestrian on the assumption that this pedestrian lacks rights not to be harmed. **Of such a case, Jon writes:**

But when our risk-imposing actions do not treat anyone else as lacking moral rights against harm—when we have an independent moral justification for the evidence-relative permission to impose harm or risk of harm on others that does not depend on this assumption—then the moral picture seems very different. Liability to defensive harm, on this view, is grounded in a particular conception of what it means to treat others with the concern and respect they are due.<sup>32</sup>

I think this is a deep and important difference that Jon has identified between the dignitary on the one hand, and the conscientious driver on the other hand. But I don't think this distinction is sufficient, as Jon maintains, to support his view that the conscientious driver retains his non-liability to defensive harm. Rather, I think this difference simply mitigates the extent to which his dangerous activity for which he's morally responsible renders him liable to such defensive harm.

---

<sup>31</sup> Quong, "Liability to Defensive Harm," *Philosophy and Public Affairs* 40 (2012), p. 69.

<sup>32</sup> *Ibid.*, p. 69.

Jon maintains that the situation between the driver and the pedestrian is completely symmetrical in the following two respects: (i) neither has rendered himself liable to defensive harms by forfeiting his rights, but (ii) each has an agent-relative permission to defend his own life, even though this would infringe the non-forfeited right not to be killed of the other.<sup>33</sup>

Such a symmetrical approach faces the following difficulty. It cannot account for the following moral phenomena. When the car veering towards the pedestrian contains just the morally responsible driver, rather than just an innocent passenger (say a young child who cannot take control of the car, whose driver has been thrown from the car), the pedestrian has more of a justification to hurl the grenade at the car in self-defense.<sup>34</sup> The difference can't be one of agent-relative permission: such permission ought to be just as great in either case. Rather, the difference is that the driver becomes liable by virtue of his moral responsibility for driving a foreseeably potentially lethal vehicle, whereas there is no plausible account according to which the passenger could become liable.

**I think Bernard Williams's famous remarks about the agent-regret of the unfortunate lorry driver are highly pertinent here:**

The lorry driver who, through no fault of his, runs over a child, will feel differently from any spectator, even a spectator next to him in the cab, except perhaps to the extent that the spectator takes on the thought that he himself might have prevented it, an agent's thought. ...We feel sorry for the driver, but that sentiment co-exists with, indeed presupposes, that there is something special about his relation to this

---

<sup>33</sup> Seth Lazar has extensively defended the symmetry of the driver and the pedestrian in the first of these two respects. See his 'Responsibility, Risk, and Killing in Self-Defense', *Ethics* (2009).

<sup>34</sup> In fact, I think the hurling of the grenade would be unjustified when just the child is in the car: such a case is essentially the same as my Case II which I discussed in Section I above, in which a trolley with an innocent bystander stuck inside is hurtling towards you.

happening, something which cannot merely be eliminated by the consideration that it was not his fault.<sup>35</sup>

It is, I think, telling that anything remotely approaching the same degree of agent-regret would be out of place if it were experienced by the person who has caused a bomb to go off by pressing "send" on his cell phone. The difference between the reactions of the two to the killings of which they're each a cause is, I think, to be explained by the fact that the lorry driver, though blameless for the manner in which he drove, realizes that he is nevertheless morally responsible for engaging in activity that he knows to run such risks. There is no such moral responsibility in the case of the cell phone operator – at least one who has not yet been alerted, by a reading of Jeff's work, to the potentially lethal effects of pressing "send." Williams notes that a lorry driver might appropriately "act in some way which he hopes will constitute or at least symbolise some kind of recompense or restitution, and this will be an expression of his agent-regret."<sup>36</sup> I would just add, that it would also be appropriate for him, and for us, to acknowledge that his having rendered himself liable to defensive killing might well be another consequence of such morally responsible, albeit nonculpable, agency.

**End**

---

<sup>35</sup> *Moral Luck*, p. 28.

<sup>36</sup> *Ibid.*

## Appendix on Jeff's Case of the Trapped Miners<sup>37</sup>

Here is the case:

***The Trapped Miners.*** Two miners are working in a mine shaft when a[n] earthquake pitches one of them against a support beam. His body dislodges the beam, causing a collapse. The miners are then trapped, each in a separate chamber. ...They learn that rescuers will reach them in five hours. But ... the combined oxygen supply would enable them together to survive for only four hours.<sup>38</sup>

Jeff dubs the miner who dislodged the beam a *Nonresponsible Cause* of a threat. If the other miner breaches the wall that separates him from the Nonresponsible Cause and kills him, there will be enough oxygen for him to survive. If he doesn't kill him, he will die, but the Nonresponsible Cause will live. Intuitively, it is impermissible to kill the Nonresponsible Cause. Moreover, Jeff maintains that there is no morally relevant difference between a Nonresponsible Cause and a Nonresponsible Threat such as our falling man in Case III. He writes:

The only real difference between the two is a difference of timing: while it is the Nonresponsible Cause's movement *in the past* that is the innocent cause of a present threat, the Nonresponsible Threat innocently causes a present threat by his movement or presence *now*. And it is hard to see how this mere difference in timing...could *by itself* make a decisive moral difference, making it permissible to kill one who is the present cause of a present threat (that is, the Nonresponsible

---

<sup>37</sup> See n. 13 {confirm note #} above.

<sup>38</sup> *Ethics of Killing*, pp. 405-6. This is, in fact, a refined version of a similar case that appears in 'Innocent Attacker'.



Threat) when it is clearly impermissible to kill one who is the past cause of a present threat (that is, the Nonresponsible Cause).<sup>39</sup>

**A problem with Jeff's trapped miner case** is that the killing of this miner would be causally upstream the good. This is because the miner stands in the way of your securing of the lifesaving good of a supply of oxygen for yourself four to five hours hence. You need to eliminate him now in order to gain access to this oxygen later.

---

<sup>39</sup> *Ethics of Killing*, p. 406.