

I. Introduction

Consider the following two examples.

The Conscientious Driver

A careful and alert driver operating a well-maintained car is passing through a residential neighborhood on her way to the cinema. A young child playing in the front yard of her parents' house darts out into the street, just a few meters in front of the car. The alert driver sees the child, but it is impossible to stop the car in time, even though she is driving at the speed limit and even though her car's breaks are well-maintained. She strikes and kills the child.

The Mistaken Resident

A vicious psychopathic serial killer is on the loose in a small town – its residents have discovered that she kills her victims quickly and without warning. Her picture has been disseminated among the town's residents who are warned that if they encounter the killer, they should immediately take defensive action if fleeing is not possible. Unbeknownst to anyone, the serial killer has a long-lost identical twin coincidentally passing through town. She stops at a random residence one night to ask for directions. Since it is a rough neighborhood and it is unusual for the resident to receive visitors this late, and since her apartment has no available egress, she answers the door armed. Seeing what appears to be the serial killer, the resident immediately shoots her, in what she reasonably believes to be necessary and proportionate defense of her and her family.

In both cases the victims are innocent, in that they have done nothing to lose their right not to be harmed. The harm that each threat imposes on her victim is wrongful. Yet the threats are not to blame for having wrongfully killed – albeit for different reasons. The conscientious driver, by driving carefully and alertly in a well-maintained vehicle, was engaging in a permissible type of activity. And she did everything that could be fairly asked of her to reduce the risk imposed on those around her. And the mistaken resident is not culpable since she was epistemically justified (though mistaken) in her belief that she and her family were in imminent danger. If her justified beliefs were not mistaken, the harm she imposed would not have been wrongful. Thus, though she is acting impermissibly, her reasonably mistaken belief grants her a full excuse.

Suppose, now, that a third party is present just as the conscientious driver's car is about to strike the child, and just as the mistaken resident is about to shoot the serial killer's identical twin. The third party is in a position to save the child by vaporizing the car and the conscientious driver within; and she is in a position to save the serial killer's identical twin by killing the mistaken resident instead. Would the third party in both cases be permitted to do so? If so, is this because her targets are liable to defensive violence? The question can be put more generally – are *minimally responsible lethal threats* morally liable to defensively killed?

According to what I will call the "responsibility-based" account of liability, which has gained prominence in the past decade, the answer is 'yes'. I will argue that the responsibility-based accounts of liability are correct in grounding liability in a threat's *responsibility* for a harm rather than in her *culpability* for that harm. But such accounts are nonetheless mistaken in concluding that minimally responsible lethal threats are liable to be defensively killed. On the account I will offer in its place, though liability is grounded in responsibility rather than culpability, a minimally responsible lethal threat is *not* liable to be defensively killed, though it is indeed permissible to kill such a threat as a lesser evil -- *even if killing the threat does no good other than saving the life of the innocent she is wrongfully threatening*. Defenders

of the orthodox responsibility-based account will likely balk at this claim. But before I defend my heterodox version of the responsibility-based account, I will say more about the concept of liability after which I will defend responsibility-based accounts of liability in general.

II. Defense of Mere Responsibility as a Basis of Liability

A person P is liable to be harmed just in case P has done something to involuntarily forfeit her right not to be harmed in that way. That is to say, if P is liable to be harmed, she has lost that right without having disposed, sold, waived, or otherwise consented to that loss. In addition, if P is liable to a harm, she is not *wronged* by that harm, and thus would have no justified complaint against the harmer for having been harmed. *A fortiori*, P has no right of self-defense against the harm to which she has made herself liable.

On the most prominent account of moral liability, P is liable to some harm only if imposing that harm achieves a particular goal apart from the goal of harming P. Thus, for example, P is liable to be harmed only if doing so will compensate someone she has wronged in the past, or if it will prevent her from wrongfully harming another. Why should we think that liability is instrumental in this way? We can answer this question by considering, broadly, the role that rights play in our moral economy. Rights can be thought of as protecting us from being used without our consent as a means to the achievement of other people's goals. This understanding of rights provides a basis for determining how and why right can be lost. When a person violates another's right not to be used as a means, the violator herself forfeits her right not be used as a means to preventing or rectifying the harms which she was wrongfully inflicting. And this is just to say that a person who violates a right can become liable to the means required to prevent or rectify the rights-violation she has committed. Thus liability is necessarily instrumental, because its *function* is to prevent or restore a rights-violation. [Rodin, JH]

Though P is not wronged by imposing upon her the harms to which she is liable, such harms still receive some negative weight in the calculation of proportionality (though considerably less than they would receive if it were imposed on a non-liable party). So if it is possible to achieve the *goals* of harming P, but through some other means which do not harm anyone, then P is not liable to be harmed. This is tantamount to building a constraint of necessity into liability. After all, if P is not wronged by imposing the harms to which she is liable, and if there is a constraint of necessity on imposing these harms, then she cannot be liable to unnecessary harms. [McMahan, KiW]

This construal of liability also serves as a defense against the criticism that liability cannot entail the *forfeiture* of a right not to be harmed. According to this criticism, the claim that liability entails the forfeiture of a right not to be harmed illicitly overgeneralizes the conditions under which a liable party can be permissibly harmed (see for example, Judith Jarvis Thomson, *Self-Defense and Rights*, in *Rights, Restitution, and Risk: Essays in Moral Theory*, 1986). Thus, for instance, it is claimed that if P has *forfeited* her right to life she can be permissibly killed for any reason at any time, for the rest of her life. But a forfeiture theory of liability which has the necessity constraint 'built-in' can indeed explain how P can regain her right to not to be killed when she desists in her culpable aggression. The description of the right P has forfeited is complex – she has not forfeited her right not to be killed *simpliciter*, but has instead forfeited her right not to be killed as a means of preventing the harm she is wrongfully imposing. Thus if killing P does not achieve this goal, then she is not liable to be killed. [Ferzan, CA; McMahan, KiW, Tadros, D&L].

According to a responsibility-based account of liability, minimally responsible threats can be liable to defensive harm. A person bears at least minimal responsibility for a threat so long as she has made relevantly voluntary choices that *foreseeably* risked contributing to the resulting threat. The criteria for

minimal responsibility include appropriate degrees of physical and psychological self-control, as well as the capacity for rational choice. A person may be merely minimally responsible because there are circumstances (such as duress) mitigating (though not eliminating) her cognitive capacities or her degree of control over her actions. But here I will be concerned solely with threats that are merely minimally responsible as a result of having an epistemically justified thought mistaken belief that what they are doing, though risky, is morally permissible. That is, I will be concerned here with minimally responsible threats for acts which are *subjectively permissible* -- what Derek Parfit calls the "evidence-relative" sense of permission -- but which are also *objectively impermissible* -- what Parfit calls the "fact-relative" sense of permission. According to Parfit, an act "would be wrong in the fact-relative sense just when this act would be wrong in the ordinary sense if we knew all of the morally relevant facts." And an act would be "wrong in the evidence-relative sense just when this act would be wrong in the ordinary sense if we believed what the available evidence gives us decisive reasons to believe, and these beliefs were true" (Parfit 2011). What the conscientious driver does is not objectively permissible (i.e., permissible in the fact-relative sense) -- after all, she is not permitted to drive to the cinema when doing so will cause a wrongful death to an innocent bystander. But driving on that occasion *is* subjectively permissible (i.e., permissible in the evidence-relative sense) since she reason to believe that it is exceedingly unlikely that her drive on that occasion would result in the death of an innocent. On the responsibility-based account of liability, the driver's evidence-relative permission insulates her from culpability, but not liability. (McMahan, ILW)

What is the basis of the conscientious driver's liability, if she is not to blame for the objectively wrongful threat she poses? Though the conscientious driver is not culpable for the wrongful harm she poses, she is still responsible for that harm. She recognized (or was in a position to recognize) that driving on any occasion carries a small but significant risk of inflicting great harm on an innocent. The risk is small enough that the act of driving, in general, is permissible. But on this occasion the risk actualizes, threatening a bystander. Having knowingly and voluntarily engaged in an activity that imposes a small but significant risk others at risk as a side-effect of pursuing her own ends is the basis of her liability to defensive action. That is, her responsibility for the objectively wrongful harm she is inflicting grounds her liability for that harm. Thus if harming the conscientious driver is necessary to avert the lethal threat she wrongfully poses, then inflicting the defensive harm is justified on the grounds that the driver is liable to that harm. So the third party can permissibly kill the conscientious driver in other-defense of the innocent. [McMahan, KiW, SDAMIT, SD&C, tCoLtMA]

Put more generally: on the responsibility-based account of liability a person is liable to be defensively harmed to avert a threat only if a) the threat, if realized, would impose an objectively wrongful harm, b) the person is at least minimally responsible for the objectively wrongful threat, c) the defensive harm is not disproportionately severe relative to the objectively wrongful threat it averts, and d) imposing the defensive harm is necessary as a means or side-effect of averting the wrongful threat.

David Rodin has argued that if the accident is caused by an unforeseeable mechanical failure such as failed breaks or seized steering, the driver is not liable since "it would be hard to describe the strike as an action of the driver at all," he says. "The driver has now become a passenger in an out-of-control vehicle, and his situation seems comparable to cases of physical compulsion like the falling fat man." But what is relevant to liability is not whether the life-threatening event *is an act committed* by the liable party, but whether the liable party is *responsible* for that event. The two can come apart, as would be the case if I were to surreptitiously replace the blanks in an actor's gun with live rounds, in which case I would be liable to lethal defensive violence, even though the act is committed by someone else. And

unlike the falling fat man, the car accident is a foreseeable event which flows from the liable party's responsible agency, manifested in the decision to engage in a risky activity, viz., driving.

One might attempt to turn the argument for the responsibility-based account on its head. A pedestrian, one might argue, assumes a risk by having chosen to traverse near a street. Should the conscientious driver threaten the pedestrian, the latter will share some responsibility for the threatened harm. As a result it would be wrong for the pedestrian to shift the entirety of the harm to the conscientious driver by engaging in defensive violence -- or so one might argue.

If the driver's risk of striking a pedestrian is the same as the pedestrian's risk of being struck by a driver, and if neither imposing the risk of striking the pedestrian nor accepting the risk of being struck by the driver is unjustified, we might then infer that the driver is no more responsible for striking the pedestrian than the pedestrian is for being struck by the driver. And from this we might conclude that the driver is not liable to be defensively attacked. This seems to be what Ferzan argues when she says of the driver and the pedestrian that "in both instances, the only responsibility is causal, and causation is reciprocal," says Ferzan. And elsewhere she says "among reciprocal risk-takers there is no liability."

But this counterargument is easily avoided. For an agent to be liable to defensive harm on a responsibility-based account of liability, it is not enough that an agent foresees a finite possibility that a given activity might result in a wrongful harm. Rather, the risk associated with the act in question must cross a threshold. If it does so, then the agent imposing the risk can be liable to defensive harm if that risk materializes; *mutatis mutandis* for the agent assuming a risk of being harmed. Now suppose that the risk of hitting a pedestrian while driving in a residential neighborhood crosses this threshold. And suppose that the pedestrian, in having chosen to take a walk adjacent to vehicular traffic, chose to undertake a risk of being struck. But we can stipulate that her risk of being struck by a car is much smaller than the driver's risk of striking someone. And we can change the example accordingly. Suppose the driver is crossing through an Amish village which rarely sees cars, but which abounds with pedestrians. Or suppose the driver is in a third-world country, passing through a very poor village in which, again, cars rarely appear. In these cases, the risk of striking *some* pedestrian, though small, is much greater than the risk that any given pedestrian will have of being struck. This is because, in the example, there are far more pedestrian than cars. In this case the risk-taking is not reciprocal between the driver and any given pedestrian. Thus we can stipulate both that a) the risk imposed by the driver crosses the threshold at which the driver becomes liable to defensive attack should risk materialize, and that b) the risk assumed by any given pedestrian does *not* cross the threshold at which the pedestrian can be held liable for the harm of being struck by a car. The propensity of being struck by a car while taking a walk in the village is so low, that she cannot be said to have engaged in an activity that has a foreseeable risk of resulting in serious harm to her. As a result, she cannot be said to have assumed a risk of harm. The responsibility and thus the liability go solely to the driver, who did indeed choose to engage in a risky activity.

The upshot is that though minimally responsible threats are not to blame for the objectively wrongful threats they pose, there can nonetheless be a moral asymmetry between such threats and their potential victims -- an asymmetry analogous to one between *culpable* aggressors and their potential victims. In both cases, the asymmetry grounds a permission to engage in defensive violence. The chief difference between the culpable aggressor and the minimally responsible threat is the degree of responsibility they bear for the wrongful harm, where culpability just is a high degree of responsibility for a wrongful harm. The degree of defensive harm to which P is liable becomes greater as the degree to

which P is responsible for that harm increases. Thus a culpable aggressor is liable to defensive harm of a greater severity than a non-culpable threat.

On the responsibility-based account, posing an objectively wrongful threat does *not* serve as a basis of liability to self-defense if the threat is posed non-responsibly. Consider the following example first posed by McMahan:

The Cell Phone Caller

A woman's cell phone has, unbeknownst to her, been rigged so that next time she answers a call the phone will send a signal detonating a bomb killing an innocent person. She has no more reason to believe that answering a call will detonate a bomb than you have reason to believe that your cell phone is similarly rigged.

It is objectively wrong for the cell phone caller to answer her cell phone. But when she does so, her act is subjectively permissible. In this regard, she is in a position similar to that of the conscientious driver. But unlike the conscientious driver, she is *not* engaging in an activity that foreseeably has a small but significant risk of harming an innocent. Although the cell-phone caller and the conscientious driver do not *culpably* pose objectively wrongful threats, they nonetheless differ with respect to how much responsibility they bear for the objectively wrongful threats they pose. The conscientious driver is in a position to recognize that she is undertaking a risky activity, whereas the cell phone caller is not. Thus, should that risk actualize, the conscientious driver is at least minimally responsible for it, whereas the cell phone caller is not. The upshot is that the cell phone caller does not qualify as a minimally responsible threat (though neither is she an innocent threat of the same sort as those who pose objectively wrongful threats but who are not acting qua persons when they do so, such as a man whose body is flung onto an infant by a gust of wind). The driver's voluntary choice to foreseeably impose a risk on others serves as a basis of liability to defensive harm; the cell phone caller's voluntary choice does not foreseeably impose a risk on others, and thus does not serve as a basis of liability to defensive harm. [McMahan, CoLtDA]

Some theorists deny that the conscientious driver, by choosing to drive, has knowingly accepted a risk of harming a pedestrian given that this extremely low-probability event was, on that occasion, unforeseeable. Whitley Kaufman writes, "suppose, for example, that the probability of this freak event was equivalent to or even smaller than that in the cell-phone case. Does that change the conclusion?" Ferzan similarly asks, "what distinguishes the sudden onset of a psychotic breakdown from the inexplicable malfunctioning of a perfectly maintained car? How is one responsible for one risk more than the other when functioning in this society could bring about either breakdown?" But as I noted, some risks but not others cross a threshold at which imposing the risk makes the risk-taker liable to defensive attack should the risk materialize and if the risk-taker was in a position to recognize the risk she was taking. It is of course an empirical claim that the probability of a conscientious driver striking a pedestrian is higher than the probability that, for example, in answering your cell phone you will set off a bomb, or that you will spontaneously suffer a murderous psychotic breakdown. If this empirical claim is indeed mistaken then the conscientious driver would not be even a *minimally* responsible threat; thus like the cell-phone caller she would not be liable to defensive harm.

Critics of this responsibility-based account of liability have argued that it introduces a problematic 'gap' between responsibility and culpability: an individual cannot be responsible for an objectively wrongful harm without also being culpable for that harm; and even if an individual can indeed be responsible without being culpable, then she is still not *liable* for the harm in this case – or so they argue. In what

follows, I defend the responsibility-based account from this sort of criticism, before moving to what I take to be a more effective criticism of orthodox responsibility-based accounts.

III. Objections to the Responsibility-Based Account

The Argument from Culpability

Some might argue that more than minimal responsibility is necessary for a threat to be liable. In addition, the threat must be to blame for the harm she is wrongfully imposing. On this sort of account the liable party must *deserve* to be harmed by virtue of having culpably engaged in wrongdoing (though of course the harm must also be instrumental to the achievement of some distinct good, such as defense, redress, or deterrence). In this case, desert functions as a necessary condition of liability. The harm to which the threat is liable must not only be deserved, but be morally unavoidable as well, in the sense that the deserved harm must, in the words of Gardner & Tanguay-Renaud, “displace a no-lesser totality of harm than would otherwise be intentionally inflicted later, whether by the person punished or by others, and whether upon the person punished or upon others.”

But as McMahan points out, such an account is excessively restrictive. The claim that desert is a necessary condition of permissibly and intentionally inflicting harm (without infringing rights) is incompatible with the rules of fault liability in tort law, which permit us to force a wrongdoer to compensate victims whom he has unjustifiably but non-culpably harmed. If a forced compensation is a deprivation, and if the infliction of a deprivation is a harm, then the claim that culpability is a necessary condition of liability to the intentional infliction of harm is incompatible with accepted standards of corrective justice.

So any account of liability insisting that a person must deserve to be harmed in order for the wrongful threat to be liable to be harmed is excessively restrictive. Still, one might argue that the liable party must be culpable if she is to be liable to significant *defensive* violence – especially lethal defensive violence. On Ferzan’s view, “it seems that the only time that we may say a defender does the right thing is when he acts against a culpable aggressor.” On Ferzan’s view, the criterion of liability to defensive force (what she calls the “triggering conditions” to defensive force) is the wrongdoer’s culpable attempt to impose an unjust threat. This is the only sort of case in which a person can be defensively attacked without infringing or violating her rights. On her view, violence directed against those who are not culpable is unjust, even when the non-culpable party is responsible for posing a wrongful lethal threat.

But the claim that liability to *defensive* harms – but not compensatory harms – requires culpability is *ad hoc*. As David Rodin points out, “defensive rights and redress rights in particular are closely linked. Both provide a personal remedy for [a wrongfully threatened innocent] to secure his rights against infringement” (and in this respect, “both differ from liability to punitive harm”). It is unclear why risk-takers need not be culpable to be subjected to the deprivations of redress, but must be culpable to be subjected to defensive harms. Consider activities that are inherently risky, even when all reasonable precautions are taken. This includes engaging in certain recreational sports (such as hunting), undertaking certain business activities (such as transporting volatile material), and owning certain exotic pets (such as venomous snakes). Should an accident occur wrongfully harming an innocent victim, that victim would be owed compensation by the individual who engaged in the risky activity, even if she had taken all reasonable precautions at the time. That is to say, the individual who is responsible for the wrongful harm imposed on the innocent would be liable to certain deprivations as a means of redressing the harms incurred, despite that she did everything that morality could fairly ask of her. If she is liable to these deprivations after the harm has occurred, as a means of bringing the victim back to a condition as close as possible to the condition she was in prior to suffering the wrongful harm, then why can’t she be

liable to deprivations prior to the occurrence of the harm as a means of preventing the harm from occurring in the first place? It might be argued that neither bodily harms in general nor lethal harms specifically would be countenanced *ex post* as a means of compensation. But this is no surprise, for at least two reasons.

First, *ex post* compensation implies that the victim, whatever harms she has suffered, is still alive. The constraint of narrow proportionality restricts the amount of harm we can impose on a liable party as a means or side-effect of preventing wrongful harm on the innocent party. This is one reason why the restriction on *preventive* harms is less severe than the restriction on *compensatory* harms – the preventive harms, but not the compensatory harms, might be necessary to prevent the wrongdoer from killing the victim.

Second, *ex post* deprivations consisting of bodily harm typically have no compensatory value. And even if they did – as in the case of organ transplants – the restrictions against harming intentionally and opportunistically are more stringent than those against harming collaterally or eliminatively. So, for example, destroying a minimally responsible wrongdoer's liver as a side-effect of preventing him from wrongfully killing an innocent might be permissible, even if it is impermissible to surgically transplant his liver into his victim as a means of preventing his death from the injuries caused by his wrongful attack. So even when deprivations consisting of bodily harm typically do indeed have compensatory value, imposing such harms can be impermissible, even if the same harm committed collaterally or eliminatively would have been permissible *ex ante*.

To better see this, consider the following example. Suppose a person engages in a risky though permissible type of activity; an accident ensues which critically harms both the risk-taker and an innocent bystander. They both require a series of surgeries to save their lives. The risk-taker has the funds to finance her own surgeries; the bystander does not. The bystander is owed compensation for the life-threatening harm the risk-taker has wrongfully imposed on her. I take it that the bystander is liable for these harms; she can be forced to pay for the bystander's life-saving surgery, even if this means that she will not be able to afford her own. In this example, compensatory liability *ex post* does indeed result in the death of a non-culpable but responsible wrongdoer, as a side-effect of saving the life of innocent victim.

The upshot is that there are indeed reasons why the restrictions on compensatory harms are more stringent than the restrictions on preventive harms – but none of them advert to the culpability of the wrongdoer. If we are willing to say that non-culpable but responsible wrongdoers are liable to *ex post* compensatory financial deprivations necessary to save the victim's life, even if it results in the wrongdoer's death as a foreseeable side-effect, then I do not see why we cannot say that non-culpable but responsible wrongdoers are liable to lethal *preventive* harms necessary save the life of the innocent. To claim that, on the contrary, culpability is a necessary basis of liability to preventive but not compensatory harms is, again, *ad hoc*.¹

¹ Of course, as David Rodin points out, there are reasons why defensive and redress rights should be treated differently. He points out that self-defense is a morally risky activity, since those who are in a position to defend against wrongful harms are typically interested parties who have to act quickly with little time to reflect on the moral complexities of liability, and with incomplete information about the apparent threat. The decision to inflict defensive harm, then, involves taking a substantial risk of wrongdoing. There are reasons of justice, then, to defer exercising defensive rights in favor of redress rights: *ex post*, we have the opportunity to properly determine

Though I believe that the argument from culpability fails on its own, I think it finds purchase when combined with another argument against the responsibility-based account. I will argue that, on the orthodox responsibility-based account, moral luck serves as a basis of liability in two ways. The first way is not problematic, but the second way is -- I do not think that moral luck can serve as a basis of liability *to be killed* when the minimally responsible party is blameless for the materialization of a risk she justifiable or permissibly imposed. After explicating this problem, I will then develop a different, better version of the responsibility-based account.

Argument from Moral Luck

On the responsibility-based account, moral luck partly grounds a minimally responsible threat's liability to be defensively attacked. This does not mean that luck alone can determine whether a person is morally liable to be attacked. A person becomes liable to defensive harm only if she engaged in voluntary action for which had reasonably foreseeable risk of resulting in or contributing to a wrongful harm. If the risk is sufficiently slight, and if the activity is a type that is not proscribed, that the risk-taker is not to blame for the harm should it actualize. But the risk-taker is nonetheless liable to defensive harm because she chose to act in the awareness of that risk. As McMahan says of the conscientious driver, "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." And since the driver chose to impose the risk for her own benefit, it is fair that *she* rather than the pedestrian should suffer the cost of the risk should it actualize. Of course, there are many drivers who imposed the same risks that she did. "That she is liable and they are not is a matter of moral luck."

A similar principle partly grounds the liability of another sort of minimally responsible threat -- the mistaken resident. Though she acts reasonably given her epistemically justified beliefs, she nonetheless chooses to try killing someone knowing that the use of lethal force in this case is unjustified. As McMahan puts it, "when one chooses to kill another person, one renders oneself vulnerable to the possibility of mistake, and if one is in fact mistaken, even if only through bad luck, one is liable to suffer the bad consequences of one's choice." Here McMahan follows Michael Otsuka, who writes that "a morally responsible agent may be held accountable for engaging in such activity that puts the life of a potentially innocent person at risk even if she acts from the justifiable (but false) belief that this person is a villain. 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."

This is not to say, however, that the conscientious driver and the mistaken resident are equally liable. The resident, unlike the driver, *chooses* to kill a person intentionally. And this is a much greater moral risk than the one associated with driving. If moral responsibility for an objectively unjust harm is the basis of liability to defensive harm, and if the degree of wrongdoer's liability varies with her responsibility, then the mistaken resident is liable to more severe defensive action than the conscientious driver. Still, in both cases, it is the attacker's bad luck that serves as part of the ground of their liability.

whether and to what degree the apparent wrongdoer is liable. This does not show, though, why liability to *ex ante* harms imposed defensively would require culpability, but liability to *ex post* remedying deprivations do not.

Ferzan rejects Otsuka's argument that the mistaken resident, liability lies in her decision to take "a gamble" which places her "moral immunity on the line". Otsuka, *supra* note 12, at 91. "After all," Ferzan writes, "they are likewise taking a gamble if they do not defend themselves, and they are gambling with their lives. I do not see how why we should say that such actors forfeit their rights to life when they do all that morality can fairly ask of them."

There are several responses here. First, it is true that if the resident chooses not to shoot, she is gambling with her own life. Thus whether or not she shoots, she is taking a considerable risk. But there is a moral difference between gambling with one's own life and gambling with the lives of others. If the resident chooses to shoot first, then she is risking the life of someone who might turn out to be an innocent. She is, in effect, gambling with the life of a possibly innocent person in order to ensure that she survives. To choose this gamble is to shift risk away from herself and onto a possibly innocent person. If it turns out that this person is indeed innocent, then it is only fair that the aggressor bear the cost. It is true that she has done everything morality can fairly ask of her – but this is consistent with claiming that she is nonetheless morally liable.

This might seem implausible, when applied to conscientious driver who, in Ferzan's words, "behaves in a cautious and admirable way". It seems *unfair* to claim that the driver has forfeited her right to life, despite having done everything that someone in her position can be reasonably asked to do by morality. David Rodin argues along similar lines. The result that the driver is liable to defensive attack is "intuitively uncomfortable", he says. He argues that the "risk of injury which the driver imposes was not proscribed given that he had fulfilled all his obligations to minimize the risk. The pedestrian has no right not to be exposed to such a risk." The tacit premise here seems to be that fulfilling all of one's obligations to minimize a risk precludes liability should that risk materialize.

But as a general claim, this cannot be right. Suppose, through no fault of her own, a diabetic out for a walk suddenly slips into a coma which will kill her unless she is administered insulin immediately. A passerby knows that a neighbor has a supply of insulin; but the neighbor is away. So the passerby steals the necessary insulin. Though this theft is justified, the passerby is liable for the cost of compensating the owner of the insulin for the window which was shattered as the only means of breaking into the house, should the diabetic refuse to pay. This is despite the fact that the passerby acted with moral justification. The lesson here is that we can be liable to certain harms even in cases where we act with justification. Is it unfair to claim that the good Samaritan, who aided the diabetic in precisely the way morality required of her, has lost her right not to be fined in compensation? Perhaps it is. But it would be more unfair to force the proper owner of the insulin to bear the cost of the broken window. Similarly, perhaps it is unfair to claim that the driver has lost her right not to be defensively attacked after having behaved in a "cautious and admirable way". But it would be more unfair to impose the costs of what the driver knew to be a risky activity on her victim. (This *comparative* dimension of liability has been under-analyzed – I will argue later on that what grounds a permission for the pedestrian to defensively attack the driver is a combination of a liability-based justification and a lesser-evil justification).

Still, opponents will say that it seems that the conscientious driver and the resident are simply victims of bad luck – and in this respect it is unfair to treat this bad luck as a basis of liability. What has gone unnoticed, however, is that there are actually *two* ways in which moral luck determines liability on the standard responsibility-based account. First, whether an individual is liable at all can be determined in part by luck. Thus, for example, whether an attempted aggressor is liable to defensive violence can depend on whether she is successful in her attempt, which can be a matter of luck. A culpable aggressor whose gun jams is not liable to be defensively killed by her target if she now has a chance to treat, since

killing the aggressor is not necessary to avert her death. Likewise, whether the conscientious driver becomes a threat, and is thereby liable to be defensively action, is a matter of luck – the only thing differentiating the minimally responsible threat from every other conscientious driver is an unforeseen mechanical failure, or the happenstantial event of a child running out into the street.

But there is a second way in which moral luck can determine liability on the standard account. *The degree of defensive harm* to which a threat is liable depends on what sorts of defensive harms are available to the victim. Thus, for example, if the threat posed by the conscientious driver can be averted through defensive force that merely breaks her arm, then she is not liable to be *killed*. It is often happenstantial what types of defensive force are available to a potential victim. And since the types of harms to which a threat is liable is determined in part by the types of defensive force available, it follows that the types of harms to which a threat is liable is a matter of luck. This is, then, a second way in which moral luck can determine liability on the standard account.

On the standard responsibility-based account, the degree to which a threat is liable can come apart radically from the degree of responsibility which she bears for that threat. Thus the role that moral luck plays in determining what sorts of harms to which a threat is liable represents another way in which responsibility can come apart from liability on the standard account. Suppose there are two persons bearing the *same* degree of responsibility for the *same* type of wrongful harm – one can be liable to a harm as severe as death, whereas the other is liable to no more than a pinch, if it turns out that in the first case no effective defensive harm less severe than death is available.

(Note that even if the criterion of necessity were not internal to liability, it would still be permissible, on the standard account, to kill the conscientious driver provided that doing so satisfies the additional constraint of being necessary to save the pedestrian's life. This version of the standard account would yield a *range* of harms to which the threat is liable. But even this account allows moral luck to determine whether a harm to which the threat is liable can be defensively imposed on the threat, since any defensive harm greater than what is necessary to avert the wrongful harm cannot be permissibly imposed).

I believe that defenders of the standard responsibility-based account are mistaken in the claim that a minimally responsible threat, such as the conscientious driver or the mistaken resident, can be morally liable to be defensively killed; they are, in my view, mistaken since I do not think that moral luck can serve as a basis of liability to be killed when the minimally responsible party is blameless for the materialization of a risk she justifiable or permissibly imposed. In what follows, I present an alternative account of liability.

IV. The Heterodox Responsibility-Based Account

The Moralization of the Lesser Evil Justification

My account is, at its core, a responsibility-based account of liability: innocent threats have forfeited their right to not to suffer the proportionate harms necessary to prevent the wrongful threats for which they are at least minimally responsible. But on my account, however, *killing* an innocent threat is a disproportionate means to or side-effect of preventing the harms that the innocent threats wrongfully pose, even if killing the innocent threat is the only way to prevent that harm. Nonetheless, killing an innocent threat is still morally justified as a lesser evil, relative to the alternative of allowing one's self to be killed. This claim will sound untenable to proponents of responsibility-based accounts of self-defense. After all, in the examples under consideration, either the minimally responsible threat will be killed, or an innocent bystander be killed. How can the former option be the lesser evil relative to the latter

option? In both cases, one person is killed. To explain how there can be a lesser-evil justification for killing an innocent threat, it is first necessary to say more about the lesser-evil justification in general.

It can be permissible to harm a person, P, even when that P has neither waived nor forfeited her right not to be harmed, if the consequences of refraining from harming P would be significantly worse from an impartial point of view. In such a case, P retains her right not to be harmed, though it is *overridden* by competing moral reasons. This assumes that rights, as agent-centered restrictions, are not absolute. If the negative consequences of abiding by an agent-centered restriction are sufficiently weighty, then we have an all-things-considered reason to do what P has a right that we not do. The right, in this case, is *infringed* rather than violated.

Even though we act permissibly when we infringes P's right not be harmed, P is still wronged; the reactive attitudes of resentment and anger can be morally appropriate, and we are morally required to compensate P for the harm we have inflicted. This is in contrast to *liability* and *consent* as grounds for permissibly inflicting harms – in those cases, P's right is neither infringed nor violated, which is to say that P was not wronged.

A lesser-evil justification for harming an innocent is an eminently familiar, consequentialist mode of moral justification in normative ethics: the deontological constraint in favor of respecting rights can be outweighed by the negative consequences of doing so. A liability-based justification, on the other hand, is deontological in its pedigree; it acts as a 'filter' on the consequences that we are allowed to take into consideration when determining what the morally appropriate use of force is. One way to put this is that liability-based justifications are *local* whereas the lesser-evil justifications are *global*, in that the reasons and values a liability-based justification brings to bear in the moral calculus derive in some way from the rights and duties of the agents in conflict, whereas a lesser-evil justification is not similarly limited – it involves an *overall* comparison of the states of affairs that would result from the actions available to the agent.

One way that a liability-based justification differs from a lesser-evil justification is in how they assess which benefits of a harm can offset that harm in the calculation of proportionality. Benefits to uninvolved third parties are generally not included in the calculation determining whether P is liable to defensive harm. Suppose, for example, that P is a voluntary organ donor. The fact that inflicting lethal defensive violence on P will foreseeably benefit the recipients of P's organs does not factor into the determination of whether P is *liable* to defensive violence; *mutatis mutandis* should P's death indirectly harm third parties.

A lesser-evil justification, on the other hand, includes *all* the benefits and harms resulting from the infliction of violence in determining whether inflicting violence on that occasion is permissible. As a result, a lesser-evil justification presumes an *axiology* of values. We cannot compare states of affairs *ex nihilo*; there has to be some prescribed list of values which function as the basis by which we compare states of affairs. At the very least, death and suffering will be included as disvalues in any plausible axiological basis for a lesser-evil justification. But even this relatively unobjectionable claim requires qualification. *Deserved* suffering and death (if there is such a thing) inflicted on culpable wrongdoers will presumably be weighed differently from the undeserved suffering and death of innocents. It would be absurd to say that choosing the death of one hundred retired war-criminals would not be the lesser evil relative to the alternative of choosing the death of one hundred retired teachers (given that these are our only options).

We might go further by claiming that whether a candidate harm is brought about intentionally versus merely foreseeably, or through inaction versus action, are relevant to the weight it ought to receive in the lesser-evil calculation. McMahan makes this point when he says that the moral relevance of these distinctions

...together establish a strong moral presumption against doing harm intentionally. When choosing the lesser evil requires *intentionally causing* that evil, the presumption against the intentional infliction of harm must be overridden. It is because of the strength of that presumption that the harm averted must be substantially greater than the harm caused. Otherwise the presumption could not be overridden. So the greater the moral significance of these distinctions is, the greater the difference must be between the harm inflicted and the harm averted in order for the intentional infliction of harm to be justified as the lesser evil.

David Rodin makes a similar point:

Whether action is justified on lesser evil grounds will in many cases depend on more than a simple comparison of magnitudes and probabilities of outcome harm and benefit... the value of states of affairs relevant to lesser evil justification depends crucially on how harms and benefits come into being, and how they are related to human agency. It is implicit in lesser evil justification that harm inflicted on an innocent bystander is a greater evil than an equivalent harm inflicted on a person liable to that harm. Similarly, harm brought about through positive agency is a greater evil than harm brought about through negative agency, and harm intentionally inflicted plausibly counts as a greater evil than harm brought about as a foreseen but unintended side effect of justified action.

It is clear, then, that the axiology of the lesser-evil justification is *moralized*, in the sense that when we determine which of the relevant alternatives constitutes the lesser evil, we do not simply compare the total harms obtaining in these states of affairs – we also morally evaluate these harms.

In addition to the lesser-evil justification, there are various other possible grounds for permissibly inflicting violence, including considerations of desert, consent, liability, enforceable duties, etc. We have already seen that considerations of *desert* can factor into a lesser evil calculus: deserved harms are of lesser disvalue than undeserved harms. Can considerations of *consent* also factor into a lesser evil calculus? It certainly seems so. It would be absurd to say that punching one hundred people who have consented to be punched would not be the lesser evil relative to the alternative of punching one hundred people who have indeed consented to be punched (given that these are our only options). If considerations of consent and desert factor into the lesser evil justification, then why not include considerations of liability as well? After all, it seems that choosing to kill one hundred people who are morally liable to be killed is the lesser evil relative to the option of choosing to kill one hundred people who are not morally liable to be killed (given that these are our only options, and that there are no other morally relevant consequences).

One might worry that given a sufficiently moralized axiology, the lesser evil justification loses its distinctive profile by merging into homogeneity with an all-things-considered justification – which functions by adjudicating between competing first-order reasons for and against a given act. But even a lesser-evil justification with a maximally moralized axiology will not be identical to an all-things-considered justification. Rather, a lesser-evil justification is *a type* of all-things-considered justification. Specifically, it is an all-things-considered justification where the reasons in favor of transgressing a person's rights outweigh the reasons against doing so. Thus, in deciding between two supererogatory acts, I might have an all-things-considered justification for choosing one over the other – but I cannot

have a lesser-evil justification since neither of the acts causes any wrongful harm. Or suppose doing ϕ is against my prudential interests; but the moral reasons in favor of ϕ outweigh the self-interested reasons not to ϕ . In this case, I have an all-things-considered reason to ϕ . But this is not a lesser-evil justification. On the other hand, suppose the reasons to kill an innocent outweigh the reasons to respect her right not to be killed; in this case there is not only an all-things-considered reason to do so, but a lesser-evil justification as well.

In summary, a lesser-evil justification is not a first-order justification for or against the infliction of violence, on a par with other first-order justification, such as a liability-based justification or a desert-based justification. Rather, it is a consequentialist meta-evaluation of all the morally relevant factors pertaining to the options available to the agent in that context. Put differently, a lesser evil justification with a maximally moralized axiology assimilates all the other first-order grounds for permissibly inflicting violence, attaches weights to them, and adjudicates among them in consequentialist fashion. When this procedure yields a justification for committing an act which transgresses someone's rights, then we have a lesser-evil justification for that act.

The Problem of Residual Injustice

Return now to the responsibility-based account of liability. A problem with this account, as it has been developed in the literature, is that it seems to yield unjust verdicts, in that it permits imposing greater harms on innocent threats than justice would permit. The defensive harms to which the threat is liable is not concomitant with the threat's responsibility for that harm, partly because (as I've noted) moral luck serves as a basis for the severity of the defensive harm to which the threat is liable. In what follows I more closely examine how liability is distributed between a threat and a victim according to the orthodox responsibility-based account, and why the manner in which it does so is problematic.

When there is a total amount of unavoidable harm and we have to choose how to distribute it between two parties, there may be some in-principle ideally just distribution of the harm. On the orthodox responsibility-based account of liability, a liability-based justification for harm distributes unavoidable harms in a way that comes closest to the ideally just distribution. Suppose that there is a set amount of harm that must be imposed on either P or J. P bears 75% of the responsibility for this predicament whereas J bears 25%. All things being equal, the ideally just distribution of the harm would be for P to suffer 75% of it, and for J to suffer 25% of it. And since each party is at least minimally responsible for the predicament they face, P would be liable to suffer 75% of the harm, whereas J would be liable to suffer 25% of it -- if the harm were divisible. But since, as a matter of (bad) luck it turns out that the harm as whole must befall either P or J, P is liable to suffer 100% of the harm since that distribution best approximates the ideal distribution.

This is why, on the orthodox responsibility-based account of liability, the conscientious driver is liable *to be killed* despite that her responsibility for the predicament is minimal. The harm which the driver threatens cannot be divided between her and the pedestrian (or, for that matter, among all those who chose to drive under similar circumstances). The harm as a whole must be imposed on either the driver (by defensively killing her) or the pedestrian. Since the driver is at least minimally responsible for engaging in an activity that carries a small but substantial risk of inflicting harm on an innocent, she rather the pedestrian should suffer the harm; *mutatis mutandis* for the mistaken resident.

It is clear, then, that a minimal difference in responsibility can be decisive in the distribution of unavoidable harms. "Even a slight degree of responsibility for a situation in which someone must die," McMahan writes, "...can be sufficient to override the presumption against intentional killing. A degree of

responsibility or culpability that would be too slight to be of any significance if the issue were punishment can be decisive when the issue is defense.” This might seem unfair, precisely because the severity of the harm imposed on the minimally responsible threat is greater than the severity of the harm which justice would have us impose upon her if the harm were divisible. In response to this McMahan says that “what this shows is only that the concept of liability is quite different from the concept of desert. Whether one deserves to be harmed is independent of whether harm is avoidable... Liability, in contrast, arises only when harm is unavoidable and must be distributed. Even if one bears some responsibility for an unavoidable harm that cannot be divided, one may not be liable to suffer the harm if someone else is more responsible.”

Yet McMahan admits that though a minimally responsible threat such as the conscientious driver is liable to be killed, there is a “residual injustice” in killing her, since she has suffered more than her “fair share”. Killing a minimally responsible threat “leaves a moral ‘remainder’ in a way that killing a culpable threatener does not”. This is evidenced by the fact that the minimally responsible threat would be owed compensation *ex post* for the degree of harm imposed upon her exceeding the degree to which she would be liable if the harm were divisible. If the threat were culpable, however – which is to say that if she were far more responsible for the wrongful harm she has caused – then imposing the entirety of the indivisible harm on her would *not* leave a “residual injustice” or a “moral remainder”, precisely because she would be liable to suffer the entirety of the harm (and her potential victim none of it) even if the harm were divisible.

So the severity of the harm to which a minimally responsible threat is liable depends on luck – but if the minimally responsible threat ends up liable to more than the amount to which she would be liable if the harm were divisible, then she is still being treated unjustly. So in cases where a minimally responsible threat is killed or severely injured, such as in the driver or resident cases, there is a sense in which she is being treated unjustly, on the orthodox responsibility-based account. But what, then, justifies imposing harms on the minimally responsible threat that are greater than distributive justice would allow?

Put simply, though imposing the harm on the minimally responsible innocent threat is unjust, it would be *more* unjust to impose the harm on the victim. So imposing the harm on the minimally responsible innocent threat is *not* unjust, from the second-order standpoint of deciding where an unavoidable injustice should fall. But it *is* unjust from the first-order standpoint which considers solely whether the minimally responsible threat has done anything to warrant that degree of harm imposed on her.

Proponents of the orthodox responsibility-based account couch the reasons why the minimally responsible threat is liable to be killed in terms of distributive justice and fairness. In the words of McMahan, “Since the driver chose to impose the risk for reasons of her own, it is *fair* that she should suffer the cost rather than imposing it on the pedestrian” (my emphasis). Ferzan notes this as well: “Jeff McMahan justifies self-defense as a matter of fairly distributing harm: if someone must suffer a loss, the claim is that it is better that the one who culpably created the situation bears the loss. Thus, when the defender and the aggressor perceive a risk of harm, it is appropriate and just for the defender to redistribute the risk of harm to the culpable aggressor.” Robert Schopp summarizes the view that fairness is what grounds the permission to kill (though he limits the claim to culpable aggressors) “When inevitable harm must be distributed, culpably causing that inevitable harm provides a morally relevant difference between victims and aggressors, justifying victims in distributing the injury to aggressors.”

It seems to me that the best way to characterize what is going on here is to say that imposing the undeserved harm on the minimally responsible innocent threat is the lesser of two evils -- since the

alternative option is to impose an even greater undeserved harm, which is an even greater miscarriage of justice.

This suggests that the lesser evil justifications and the liability-based justification are more closely related than has been realized. On the account I propose, a minimally responsible threat is liable only to the amount of harm concomitant with the degree of responsibility she bears for the threat that the imposed harm prevents. Anything above that amount (what McMahan calls the “moral remainder”) is justified only as a lesser evil, relative to the alternative of imposing the harm on the other party. Return to the example of the fixed harm which must befall either P or J, where P bears 75% of the responsibility for this predicament and J bears 25%. On the standard responsibility-based account of liability, P is morally liable to suffer 100% of the harm and J is morally liable to suffer none of it, provided that the harm cannot be divided between them. But on my account, P is morally liable to suffer only 75% of the harm – but it is permissible to impose the remaining 10% on P as the lesser evil, since the alternative is to impose 100% of the harm on J.

If the harm were divisible, then what the minimally responsible threat would be liable to would remain unchanged – each would still be liable to the amount concomitant with the degree of responsibility she bears: 75% and 25% of the harm respectively. This shows that on my modified account, *moral luck no longer determines the severity of the harm to which one is liable*. On the standard responsibility-based account of liability, the degree to which I am liable for a harm depends on the sorts of defensive action available to the potential victim – and thus can come apart from the amount of harm that would be delegated by the amount of responsibility I bear for the predicament.

Of course, on my account, what counts as the lesser evil still depends on moral luck. Whether defensive harm can be permissibly imposed upon the minimally responsible threat depends in part on whether it so constitutes the lesser evil. And this, in turn, depends on comparing the option of imposing the defensive harm which happens to be available to the defender of the potential victim against the alternative of allowing the potential victim to absorb the harm posed by the minimally responsible threat. But moral luck as a basis for the permissible infliction of violence finds a more natural home in the lesser evil justification than in liability-based justifications. This is because reasons to avoid the lesser evil are compatible with rights-infringement. That is, if I have a lesser-evil justification for harming someone, this is consistent with the claim that the victim has done nothing to lose her right not to be harmed. So to claim that moral luck can ground a lesser-evil justification is far less objectionable than the claim that moral luck can ground a liability-based justification. The claim that moral luck can be a factor in *losing* one’s right not to be killed is stronger than the claim that moral luck can be a factor in having such a right *overridden* when necessary to avert a significantly greater evil.

Those who are inclined to deny that moral luck can be grounds for losing one’s right would presumably be more amenable to the claim that moral luck can be grounds for having one’s right overridden. Suppose a villain kidnaps 21 innocents – she picks 20 of them at random and ties them to the right trolley track; the remaining one she ties to the left. A trolley is headed towards the 20. We can permissibly divert the trolley towards the one innocent to save 20, even though the one innocent did nothing to warrant being killed – it is just a matter of luck that she ended up being there. Her misfortune is no grounds for claiming that she cannot be permissibly killed in order to avert a greater evil – though it is grounds for claiming her right not to be killed is merely overridden and not forfeited.

The upshot is that an advantage of the account I am proposing over the standard account is that moral luck no longer determines the degree of defensive harm to which a morally innocent threat is *liable*. If

this account is correct, then proportionate and necessary self-defense against a minimally responsible threat is justified neither by a liability-based justification alone, nor by a lesser-evil justification alone, but by both. Thus responsibility-based *and* desert-based accounts of self-defense are incorrect, since they both maintain, in Ferzan's words, that "self-defense is not about committing the greater or lesser evil. It is about giving the defender a right to act on the basis of the threat that she perceives." If my account is correct, then self-defense against minimally responsible threats is indeed about committing the greater or lesser evil, in addition to giving the defender a right to act on the basis of a wrongful threat.

On the modified account, the liability a wrongful threat bears is determined in part by a function which takes as its variables the severity of the moral wrong, and the degree of responsibility the agent bears for that wrong. And the degree of responsibility she bears is determined in part by the antecedent probability that the harm will occur. The crucial difference between the modified account and the standard account is how the *degree of responsibility* determines an agent's liability.

On the standard account, an agent who bears a very low degree of responsibility for a wrongful death (such as the conscientious driver) can nonetheless bear a very high degree of liability, in that she can be liable to be killed, if necessary, to avoid that wrongful death. Thus, on the standard account, the degree of liability that a wrongful threat bears is not necessarily concomitant with the degree of responsibility she bears. Of course, this certainly doesn't mean that the degree of responsibility is *irrelevant* to the degree of liability, on the standard account – but on the standard account liability can outstrip responsibility in a way that some of have found problematic.

On the modified account, the degree to which an individual is liable for a wrongful harm is more tightly constrained by the degree to which she is responsible for that harm. The driver, for example, bears some responsibility for killing the pedestrian despite that she was driving safely – thus she bears some liability as well, though she is not liable to be defensively *killed* on the modified account. For example, she might be liable to have her arm broken, if that is the only way to save the pedestrian's life. (This might be justified as the lesser evil as well, though if this were the only basis for imposing the harm on the driver then she would be entitled to compensation for infringing her right not to have her arm broken). Of course, we can disagree on precisely what harms the driver is morally liable to suffer, while agreeing that it is a harm substantially less than death, on the grounds that the degree of *responsibility* she bears for the threat she ends up posing is small.

It is important to note that if the innocent threat is not liable *at all* for the wrongful threat she poses (as is the case for the cell-phone caller), then killing the innocent threat is *not* the lesser evil, even if it the only way to prevent the lethal threat she poses. After all, in this situation, we have two morally symmetrical non-responsible parties. Given that there are no other morally relevant consequences of killing the innocent threat or allowing the innocent threat to kill, there are no grounds for claiming that killing the innocent is the lesser evil. (Though there might be agent-relative reasons for the victim to prevent the lethal threat by killing the innocent threat).

Return, then to the original example. A child darts out in front of the path of the conscientious driver – though the driver is driving safely, there is no way for her to stop the car in time. A third party, armed with a vaporizer, can vaporize the car and the driver with it. There is no other way available to save the child's life. On my account, the conscientious driver is *not* liable to be *killed*, though she is liable to some lesser harm, since driving is an activity that foreseeably poses risks to pedestrians and other bystanders.

Because we have only the option of killing the liable driver or allowing the non-liable innocent to be killed, killing the driver is the lesser evil.

This raises the question: in *killing* the driver, are her rights infringed? Yes, since she was only liable to a harm less than that of death. As a very rough analogy, suppose that a debtor owes a creditor \$75 which he has culpably failed to pay back. The creditor is entitled to retrieve the money she is owed by means that would otherwise count as theft. (Suppose they are in a state of nature, and there are no authorities to which the creditor can address her grievance). The debtor is about to leave the country and disappear forever. The creditor has a chance to surreptitiously retrieve her money from his wallet – and it is her last chance. But the debtor has only a 100 dollar bill – and there is no time for the creditor to pay her back the difference. I believe it is permissible for the creditor to take the 100 dollar bill, though the debtor is liable only to the loss of only 75% of that amount. Taking the remainder can be justified as the lesser evil, since the only relevant alternative is for the creditor to suffer the loss of the 75 dollars to which she is entitled, which is worse than it is for the debtor to suffer the loss of the 25 dollars to which he is entitled – partly because the debtor but not the creditor is responsible for this predicament in the first place.

(Note that for the creditor to take the \$100 bill to give it to someone *else* owed \$75 by a *different* debtor would not be justified by considerations of avoiding the greater evil, since the first debtor does not owe the second creditor anything – thus *all* the money taken from the first debtor would constitute a violation of her rights, which is too egregious a violation to be outweighed by the good done by forcibly redistributing wealth from her to someone who is owed money by someone else).

Likewise, though defensively killing the conscientious driver infringers her right not to be *killed*, doing so is permissible despite that it violates the constraint of narrow proportionality, since the alternative would be substantially worse. Like the debtor, the driver had more harm imposed upon her than to which she was liable. But this was permissible, since it would have been worse to impose all of that harm on the party who bears no responsibility for the predicament that they are in.

The main points of the heterodox responsibility-based account I have developed can be generalized, as follows. Defensively killing a minimally responsible threat as a necessary means or side-effect of averting a lethal harm she would otherwise inflict on a non-responsible party violates the narrow proportionality constraint. This is because the severity of the defensive harm to which a threat is *actually* liable is determined by the severity of the defensive harm to which she *would* be liable if the harm could be divided between her and her potential victim. Since a minimally responsible threat, by definition, does not bear enough responsibility to be liable to lethal defensive harm in such a case, she is not actually liable to a lethal defensive harm. However, there is still an agent-neutral justification for killing the minimally responsible threat: doing otherwise would allow a lethal harm to befall a non-responsible party, which is substantially worse from an impersonal point of view. The benefit of avoiding this outcome permits infringing the rights of the minimally responsible threat by killing her. Before exploring the implications of this account, I will defend it against two criticisms.

VI. Objections to the Heterodox Responsibility-Based Account

Objection 1: Lesser Evil and the Restriction on Intentionally Killing

One might raise the following sort of objection to my account. Killing someone who is not liable to be killed can be justified by reasons to avoid a greater evil only if the evil avoided is very significant. This is because killing someone who is not liable to be killed is a very serious rights-transgression. For there to be a lesser-evil justification for killing a non-liable individual, it is not enough that it would save two

lives, or three. Rather, many lives must be saved. Seeing as how the restriction against killing a non-liable innocent is quite strong, why believe that there is a lesser-evil justification for killing a minimally responsible threat? After all, if the evil of allowing *two* innocents to be killed is not strong enough to provide a lesser-evil justification for killing a single moral innocent, then why believe that the evil of averting a harm to *one* non-liable innocent – (such as the child in the conscientious driver case) – is strong enough to provide a lesser-evil justification for killing a *single* somewhat liable innocent threat – (such as the conscientious driver)? That is, can considerations of distributive justice – namely reasons to shift the cost of being killed from the non-liable party to the somewhat-liable party – provide a reason strong enough to override the restriction against killing someone who is not liable *to be killed*?

The answer to this depends on how liable the minimally responsible threat is. If I am liable to no more than a broken finger, then it is unlikely that reasons to avoid a greater evil are strong enough to override the restriction against killing the liable party intentionally. It would indeed be worse for someone who is liable to no harms at all to be killed than it would be for someone who is liable to a broken finger – but in my view, it is not *so* much worse that the reason to avoid this greater evil overrides the restriction against killing someone who is not liable to be killed. So if the degree of liability that an innocent threat bears is vanishingly small, then it is very unlikely that she can be permissibly killed as the lesser evil. But neither is it the case that she must be liable *to be killed* in order for it to be permissible to kill her as a means or side-effect of saving the life she is wrongfully threatening. An individual must bear more than a modicum of liability in order for a lesser-evil justification to apply, for reasons similar as to why an individual must bear more than a modicum of responsibility in order for her to be liable to be killed. As McMahan puts it, “the idea that responsibility can ground liability only if it exceeds a certain threshold is, in theoretical terms, neither ad hoc nor arbitrary. If there is a standing moral presumption against the killing of a person, and in particular against the intentional killing of a person, it should be unsurprising if it takes more than a modicum of responsibility to override it.”

The heterodox responsibility-based account I have developed has the effect of lowering the bar with respect to how much liability an individual must bear in order to be permissibly killed in self-defense. Thus a defender of the responsibility-based account of liability need not argue that the conscientious driver or the resident is liable *to be killed* in order to show that, in those circumstances, they can be permissibly killed. But if a very small degree of liability is not enough to ground a lesser-evil justification for killing the threat, then doesn't this mean that my account, though it *lowers* the bar with respect to how much liability an individual must bear in order to be permissibly killed in self-defense, makes the application of permissible self-defense more stringent? I do not think so; on my view, a minimally responsible threat for a *lethal* harm will always bear substantial degree of liability for that harm, even if it doesn't rise to the level of being liability to be killed. But others might demur, in which case the application of my account will indeed result in more stringent conditions for imposing self-defense against minimally responsible threats.

Objection 2: The Non-Fungibility of Liable Threats

If reasons to avoid the lesser evil are partly what explains why some (or all) minimally responsible lethal threats can be permissibly killed, then it might seem that minimally responsible lethal threats are *fungible* as targets of permissible violence. The following example will help explain what I mean by this.

Suppose P1 is a pedestrian threatened by conscientious driver D1. Elsewhere, P2 is a pedestrian threatened by conscientious driver D2. P1 has a vaporizer which she can use to defend herself – P2 does not. So P1 but not P2 has a chance to save her own life. Now suppose that, in addition to a vaporizer, P1 has another advanced bit of technology: a transporter. She can use it in only one way: to beam D1 into

D2's car, while beaming D2 into D1's car. One might think that if reasons to avoid the greater evil is partly what explain P1's permission to use her vaporizer, then there is nothing preventing her from using the transporter as well, even if just on a whim. After all, D1 is no more or less liable than D2. So if killing D1 crosses the threshold of liability at which killing her is the lesser evil relative to the alternative of allowing P1 to die, then D2 also crosses the threshold of liability at which killing her is also the lesser evil relative to the alternative of allowing P1 to die. So there seems to be nothing militating in favor of killing D1 rather than D2. It is in this respect that the liable targets might be thought to be fungible on the heterodox responsibility-based account. And to the extent that their fungibility seems problematic, it is a problem for my account.

This fungibility is a product of how the lesser evil justification functions. Recall that the lesser evil justification is *global*, insofar as it evaluates the relevant states of affair *en toto*, and prescribes action according to an axiological comparison of these states of affairs. It does not have the resources to delimit the reasons and values it brings to bear in the moral calculus to those that derive in some way from the rights and duties of the agents in the localized conflict. This is why reasons to avoid the greater evil seem indifferent between the sacrifice of D1 and D2 – it is enough that they are equally liable, and that the sacrifice of any one of them will prevent the death of a non-liable innocent.

Recall, however, that on my account, reasons to avoid the greater evil act in tandem with the liability of the minimally responsible threat. And recall the formal structure of liability: P is liable to a harm only if that harm is necessary to avoid the harm *for which P is responsible*. This is why, in general, one wrongdoer cannot be harmed as a necessary means of preventing a qualitatively identical wrongful harm threatened by another unrelated wrongdoer -- even if they are both equally liable for these wrongful harms. Since D2 is not liable for the threat that D1 poses (and vice versa), for P1 to kill D2 instead of D1 would not be the lesser evil. In fact, it would make things worse. By activating the transporter and using the vaporizer, P1 will have substituted a state of affairs in which P2 and D1 are killed, with one in which P2 and D2 are killed. Though the total number of deaths is the same, D2 (unlike D1) is *not* liable for the threat that P1 faces. Thus to kill D2 would be to kill a morally innocent threat. This can be permissible, if doing so averts a substantial harm, but by hypothesis only the life of P1 is saved.

The upshot is that the account I have outlined does not illicitly permit inter-substituting equally liable threats, despite that part of what grounds the justification for killing a minimally responsible threat is that doing so would avert a greater evil.

IV. Advantages and Implications of the Heterodox Responsibility-Based Account

The difference between the orthodox and the heterodox versions of the responsibility-based account might seem small – but it has significant implications.

Implication 1: Multiple Minimally Responsible Threats

Suppose that you are being wrongfully threatened not by a single minimally responsible threat (such as the conscientious driver or the mistaken resident) but by many of them – dozens or hundreds. On the standard responsibility-based account of liability, each is morally liable to be defensively killed. It seems, then, that the number of minimally responsible threats is irrelevant to the liability of each threat, and thus the permission to kill each threat. There is no point at which killing these threats violates the proportionality constraint, since each threat is liable to be killed. Yet it seems implausible to say that I can kill an indefinite number of conscientious drivers so long as it is necessary to save the life of their potential victim. This is a problem for the standard account.

McMahan argues that the standard responsibility account can avoid this problem. Recall his claim that the killing a minimally responsible threat leaves a “moral remainder” or a “residual injustice”. When multiple minimally responsible threats are killed, these remainders aggregate. “Given sufficient numbers,” McMahan says “they may eventually outweigh the defender’s claim to priority.” This is an ingenuous solution, but I think it is unavailable to a defender of the standard account. McMahan seems to want things both ways – he argues that the minimally responsible threat is liable *to be killed* -- and yet argues there is a residual injustice in killing the threat. But I will argue that this notion of a moral remainder is unavailable to proponents of the orthodox account.

Recall the two roles that moral luck plays in the standard account: it partly determines *whether* the minimally responsible threat is liable, and it partly determines *the degree to which* the minimally responsible threat is liable. Since it so happens that there is no way to distribute the defensive harm in a way that ideal justice would require of us, the entirety of the harm falls on the minimally responsible threat – she is unlucky. If, *ex post*, an opportunity arises to compensate the minimally responsible threat, effectively re-distribute the burden between them in a way that more closely follows each party’s degree of responsibility for the initial harm, then the defender is required to do so. If compensation is not possible, it is just bad luck for the minimally responsible threat. In this case, how do we make sense of the supposed residual injustice? After all, the threat’s rights are not being violated or infringed. She has no basis for complaint against the defender. She has been treated by the defender in precisely the way that she ought to be treated – both from a fact-relative and an evidence-relative standpoint. By hypothesis, a saint would have acted in precisely the way that the defender did.

It is true that in a more just world the defensive harm would have been divided more equitably. But once we admit that moral luck serves as a part of the basis of the minimally responsible threat’s liability, there are no grounds for claiming that there is a residual injustice when we kill the minimally responsible threat. Perhaps there is a cosmic injustice, in that the minimally responsible threat has a complaint against the gods or fate that things turned out the way they did. But once we allow moral luck to serve as a basis of liability, the complaint against fate cannot serve as a basis for thinking that there is a residual injustice of the sort that can influence a permission to defensively harm the minimally liable threat. To claim that moral luck serves as a basis of liability *just is* to claim that whatever sense in which it is a departure from justice to kill the minimally responsible threat is not a sense relevant to determining whether it is permissible to kill that threat. To claim otherwise is an attempt to have one’s cake and eat it too.

Thus I do not think that the standard account has the resources to explain the sense in which there *is* intuitively, a residual injustice in the killing of a minimally responsible threat – where the residual injustice is additive in such a way as to provide grounds for thinking that killing many minimally responsible threats is unjust. But my modified account deals with this problem easily, partly by eliminating moral luck as a basis of the degree of harm to which the minimally responsible threat is liable. On the modified account, killing a minimally responsible is permissible in part because doing so is the lesser evil. And as I explained, a lesser-evil justification for inflicting harm is rights-infringing. Thus the minimally responsible threat does indeed have a basis of complaint – she was harmed in a way that exceeded the degree of harm to which she was liable. This explains the *ex post* requirement to compensate the victim. And it shows why killing *many* minimally responsible threats is wrong – if the rights-infringements is additive, there comes a point where killing the minimally responsible threats simply no longer counts as the lesser evil. Indeed, I tend to think that the pedestrian is not permitted to save her own life if the only way to do so is kill *two* conscientious drivers. The upshot is that I think the

modified account, unlike the standard account, yields the intuitively correct result that it is impermissible to kill an indefinite number of minimally responsible threats.

Implication 2: Re-Emergent Moral Symmetry?

Suppose I have an all-things-considered moral justification for infringing someone's right not to be harmed as a means or side-effect of averting a greater evil. Can the victim engage in defensive harm against me in order to prevent me from infringing her right not to be harmed? It might seem that a justification for infringing a right precludes liability to defensive action. But as we have seen, there are cases in which justification is compatible with liability. Recall the example of the good Samaritan breaking into a house in order to obtain insulin necessary to save the life of an innocent who has suddenly slipped into a diabetic coma. The owner of the house would be owed compensation by the good Samaritan, even though she did everything that morality requires of her. It seems, then, that one can lose rights as a result of doing what morality requires you to do. Indeed, many have argued that infringing an innocent's right not to be killed can serve as a basis of liability to defensive harm. On their view, though killing the innocent as a side-effect is justified, to ask the innocent to sacrifice herself by refraining from engaging in defensive harm is to ask too much of them (unless, of course, the good that will result is enormous). On this view, the threshold at which it is permissible to kill innocents as a side-effect of promoting the greater good is *lower* than the threshold at which innocents are required to sacrifice themselves. Thus, for example, the civilian victims of a tactical bomber can engage in defensive violence against the bomber, even if the bomber has a lesser-evil justification for killing the civilians.

I will not take a side on this issue here. But if it turns out that the lesser-evil justification does *not* defeat liability to defensive harm – if, for example, the tactical bomber is indeed liable to defensive harm despite her lesser-evil justification – then the implications in the ethics of self-defense are considerable, on the modified account I have developed. This is because, on the modified account, the permission to defensively kill minimally responsible threats is grounded partly in reasons to avoid a greater evil – namely, the evil of imposing a lethal harm on a non-liable party. If innocent civilians cannot be reasonably asked to sacrifice themselves for the greater good by refraining from engaging in lethal defensive force against the tactical bomber, then by the same reasoning, the conscientious driver cannot be reasonably asked to sacrifice herself for the greater good by refraining from engaging in lethal defensive force against our defensive attack.

There are, of course, significant differences between the conscientious driver and the innocent civilians. The driver, unlike the civilians, posed a wrongful threat necessitating the pedestrian's attempt to engage in lethal self-defense. And the driver, unlike the civilians, is minimally responsible for that threat. Though the driver is minimally responsible for posing the wrongful threat which necessitated defensive violence in the first place, and though this functions as a basis of the driver's liability, she is still not liable *to be killed*. Thus the driver's rights, as I have argued, are indeed infringed when the pedestrian rightfully attempts to shift the cost of the impending harm onto the driver. And if justified infringements of the right against being killed serve as a basis for liability to defensive harm, then the driver can engage in defensive counter-attack against the pedestrian, despite that the pedestrian acts with all-things-considered justification.

The upshot is that if the lesser-evil justification does *not* defeat liability to defensive harm, then there can indeed be a moral symmetry between minimally responsible threats and their potential victims – both are permitted to kill the other, though for different reasons. The potential victim of the minimally responsible threat is permitted to kill in self-defense since doing so is the lesser evil, whereas the minimally responsible threat is permitted to kill in self-defense because she is being wronged. This

stands in stark contrast to the standard responsibility-based account, according to which the relationship between a minimally responsible threat and her potential victim is morally asymmetrical, in that only one of the two has a moral permission to engage in lethal self-defense.

The consequences of this re-emergent moral symmetry are far-reaching. Consider, for example, the liability of combatants in war – specifically, those who are epistemically justified but mistaken in their belief that the war in which they are fighting is just. They are in a position not morally unlike that of the mistaken resident. Such combatants, because they pose threats that are wrongful in the fact-relative sense, are liable to defensive harm. Though they are not *to blame* for the wrongful threats they pose, they still count as minimally responsible threats and are thus liable to be attacked, since they are in a position to recognize that intentionally killing someone is always a morally risky act. On the standard responsibility-based account, epistemically justified mistaken threats are liable to be defensively killed as a necessary means to preventing the wrongful harms for which they are minimally responsible. But on the responsibility-based account I have developed, the degree of liability they bear is proportionate to their responsibility for the threat they are wrongfully posing. And though the wrongful threat they are posing is very great – viz., unjust killing – their responsibility is minimal, since they are by hypothesis epistemically justified in thinking that the threat they pose is not wrongful. Killing them is permissible nonetheless, since it is the lesser evil relative to the alternative of allowing their non-liable targets (whether they are civilians or combatants) to be killed. But if a lesser-evil justification does *not* defeat liability to defensive harm, then these epistemically justified mistaken combatants can engage in permissible self-defense, despite that they are fighting in an unjust war.

This account does not reintroduce wholesale the moral equivalence of combatants, according to which combatants on all sides in a war have a moral permission to target one another. This is because combatants are not always epistemically justified in thinking that the war in which they are fighting is just. But it does mean that moral equivalence of combatants applies more often than it has been thought, given a liability-based analysis of the morality of war.

One might argue that even if the heterodox account I have developed here is correct and even if justification does *not* defeat liability to defensive violence, epistemically justified combatants on the unjust side in a war cannot permissibly engage in self-defense. This is because by doing so they are not just killing those who are not liable to be killed, but are also promoting the unjust aims of the war in which they are fighting, such as the annexation of foreign territory, political subjugation, etc. But this is unconvincing. The degree to which any given combatant promotes the unjust aims of the war in which she fights is very small – it is unlikely that this small contribution is enough to override or outweigh her permission to engage in self-defense against lethal threats to which she is not liable.

The upshot is that if justification does *not* defeat liability to defensive violence, then the modified account complicates the conditions under which it is permissible to engage in self-defense in war. There will be wars, or more likely, engagements within wars, in which combatants on the unjust side are permitted to engage in lethal self-defense despite that the cause which they are promoting is unjust.

V. Conclusion

The modified account I have developed has three key advantages over the standard account. First, it is able to explain why the *number* of minimally liable threats is relevant to the permission to kill them. Second, moral luck no longer determines the degree of defensive harm to which a threat is liable. Third, liability no longer radically outstrips responsibility, as it does on the standard account. It preserves the intuition that the conscientious driver or the resident has not forfeited their right against being *killed*

while still permitting them to be killed – all within an agent-neutral framework for self-defense. For these reasons, I think we ought to adopt in the modified account in place of the standard account. Doing so, however, has important implication in the morality of war, should it turn out that justification defeats liability. Specifically, it allows combatants who are epistemically justified but mistaken in the belief that the cause for which they are fighting is unjust, to engage in lethal self-defense against those who are attacking them.