

Rights-Based Justifications for Self-Defense: Defending a Modified Unjust Threat Account

Shannon Brandt Ford
Curtin University

ABSTRACT: I defend a modified rights-based unjust threat account for morally justified killing in self-defense. Rights-based moral justifications for killing in self-defense presume that human beings have a right to defend themselves from unjust threats. An unjust threat account of self-defense says that this right is derived from an agent's moral obligation to not pose a deadly threat to the defender. The failure to keep this moral obligation creates the moral asymmetry necessary to justify a defender killing the unjust threat in self-defense. I argue that the other rights-based approaches explored here are unfair to the defender because they require her to prove moral fault in the threat. But then I suggest that the unjust threat account should be modified so that where the threat is non-culpable or only partially culpable, the defender should seek to share the cost and risk with the threat in order for both parties to survive.

KEYWORDS: Self-defense, moral justification, culpability, human rights

1. INTRODUCTION

Meg tries to shoot Sam in order to kill him and take his wallet. If Sam is fortunate enough, in wrestling with Meg, to turn the gun onto her, and thereby kill Meg, then a reasonable person should conclude that, all other things being equal, Sam was morally justified in killing Meg. The defender (Sam) is permitted to kill the attacker (Meg) because he is entitled to protect himself from a threat that endangers his life unjustly. This is a straightforward example of justified self-defense, which says that in situations where an attack is likely to be lethal or seriously harmful, a victim is morally permitted to kill an unjust attacker in order to protect themselves. Rights-based moral justifications for killing in self-defense—a common source for morally justifying the use of lethal force¹—seek to explain how a person who is unjustly attacked has the right to take drastic action to defend her life. A key in such cases is establishing a moral difference between

the defender and the attacker. In other words, what establishes the moral asymmetry between the life of the attacker and that of the defender?

In her seminal article, Judith Jarvis Thomson argues that killing in self-defense is morally justified when one person is an immediate deadly threat to another person without a sufficiently just reason.² The attacker lacks a proper moral justification for endangering the defender's life and this is sufficient to establish moral asymmetry between the defender and the unjust threat. Here I refer to this as the Unjust Threat Account for justified killing in self-defense, which I go on to examine in some detail below. Although Thomson's account has been important to the self-defense debate, it is by no means the predominant view. Those who are uncomfortable with the notion that merely threatening another's life is sufficient to permit killing in self-defense want to set the threshold for moral asymmetry much higher. Seumas Miller, for example, requires that a person be morally culpable for posing an immediate deadly threat to the defender. That is, the attacker intends to cause serious harm to the defender.³ I refer to this view as the Culpable Threat Account. I also examine a third approach to the question of moral asymmetry, which has recently gained traction in the self-defense debate. Jeff McMahan's Responsible Threat Account does not require that an immediate deadly threat intends harm to the defender. But he insists that the threat is sufficiently liable for endangering the defender.⁴ In comparing these three rights-based accounts for justified killing in self-defense, I demonstrate that the important difference is the additional conditions required by the culpable threat account and the responsible threat account respectively. The culpable threat account requires that the threat is culpable for her attack on the defender (*culpability condition*). Likewise, the responsible threat account says that the threat must be sufficiently liable for her threat to the defender's life (*liability condition*).

I argue that these additional conditions are not necessary in cases of strict forced choice between lives. I demonstrate that the unjust threat account is sufficient with its four basic conditions. That is, killing in self-defense is sufficient when: 1) Meg is an immediate deadly threat to Sam (*immediate threat condition*); 2) Sam's primary intention is to preserve his own life, which can be achieved by killing Meg (*defensive intent condition*); 3) Sam's only reasonable option for preventing being killed is to kill Meg (*necessity condition*); 4) Meg does not have a sufficiently just reason for threatening Sam's life (*unjustified condition*). But then I demonstrate that the use of strict forced choice between lives scenarios are misleading for grounding morally justified killing in self-defense. This, I suggest, is because they are less likely to occur than a case where the defender has some leeway to accept risk and cost. Next I argue that the unjust threat account should be modified to incorporate calculations of risk and cost to both the threat and defender based on the degree of liability to defensive harm. Importantly, I argue that the threat's culpability plays a role if there is any leeway for the defender to take on cost and risk. In contrast, Thomson's unjust threat account makes a mistake in attributing no role to the culpability of an unjust threat. I conclude that the defender is obliged to share the cost and/or risk of harm *equally* from a non-culpable threat in order for both parties to survive (*shared costs and/or risks*

condition). But increasing the culpability of the threat weakens the defender's obligation (*variable culpability condition*).

2. ESTABLISHING MORAL ASYMMETRY

A rights-based moral justification for killing in self-defense is grounded in the idea that a victim of aggression has a basic right not to be killed. That is, all human beings have a right not to be killed by an unjust aggressor. David Rodin, for example, argues that a coherent explanatory account of self-defense can be constructed around the idea of personal rights.⁵ The protection of the innocent defender's life is the goal of such an act of self-defense. But if the justification of self-defense is based on the right not to be killed then it seemingly fails because a human being (the attacker) is killed. Therefore, suggests Fiona Leverick, a satisfactory explanation for justifying killing in self-defense based on a right not to be killed must demonstrate why the defender's life can be preferred to the life of the attacker.⁶ In other words, what establishes the moral asymmetry between the life of the attacker and that of the defender? Generally, this is done by suggesting that the attacker somehow forfeits his right not to be killed when he threatens another person's life. Suzanne Uniacke, for instance, argues that the use of force in self-defense does not violate its victim's right not to be killed since, as individuals, we possess this right only insofar as we are not "an unjust immediate threat to another person's life or proportionate interest."⁷ Miller argues that rights forfeiture should be based on the notion of a suspended right. This is when a right is suspended under certain conditions but not permanently cancelled.⁸ Rights forfeiture is made plausible by this notion of a suspended right. So it is permissible to kill an attacker in self-defense when he has temporarily forfeited his own life because he is threatening the life of the defender.

One approach to justifying killing in self-defense is what I am calling the *unjust threat account*. This says that the necessity of killing in self-defense is grounded in the act of repelling or warding off an unjust immediate deadly threat. The right to kill to defend a life derives from the defender's danger, from the fact that the defender's life is unjustly threatened.⁹ A well-known example of the unjust threat justification for killing in self-defense is provided by Thomson. She concludes that an unjust threat lacks a right not to be killed when he is about to violate a defender's right not to be killed and this can be prevented only by killing him.¹⁰ In a case she calls *Villainous Aggressor*, Thomson describes a scenario where a man in a truck is deliberately trying to run you down and the only way you can save yourself is by blowing up the truck. She argues that it is *morally justified*, not merely *excusable*, for you to blow up the truck and kill the driver in defense of your life.¹¹ That is, killing the driver becomes the right thing to do rather than an action that is wrong but for which you are not entirely to blame. In a second case that Thomson refers to as *Innocent Aggressor*, the driver of the truck is still trying to run you down but this time he is entirely without fault for what he is doing. Perhaps, she suggests, someone has injected him with a drug that makes him go temporarily insane.¹² It is not his fault that he is going to kill you but he will kill you if you do not blow up the truck. Thomson argues that killing in self-defense is

justified in this case just as it was in the case of *Villainous Aggressor*. The attacker's fault has no bearing on your right to defend yourself.¹³ Thomson's third case describes a scenario based on an *Innocent Threat*. In it, she suggests that a fat man, who has been pushed off a cliff, will fall on you and kill you unless you shift the position of an awning. But if you do this, the falling man will be deflected onto the road below and he will die.¹⁴ Although the falling man has done nothing at all to contribute to the deadly threat to you, Thomson argues that it is permissible to shift the awning to defend yourself, even though you know that your actions will kill the falling man.¹⁵ Thomson refers to these three cases—where it is permissible for you to kill a person in defense of your life—as “Yes” cases. In contrast, Thomson's “No” cases describe three scenarios where a person uses a bystander to defend themselves.¹⁶

Thomson's unjust threat account requires four conditions for morally justified killing in self-defense. First, the threat must be an immediate danger to another person's life. The danger is both likely to occur and of a seriously harmful nature. That is, the subject of the danger is likely to either be killed or seriously injured by the unjust threat. If Sam has the capability to defend his life by killing Meg first, then he only has moments to make the decision and act. The *immediate threat* condition permits the defender to ward off an unjust deadly threat to the defender's life in the moment that it clearly presents itself. Second, the defender's primary intention is *defensive*. That is, the defender's goal is to preserve his own life and this can be achieved by killing the threat. Third, as argued by Joanna Firth and Jonathan Quong, one cannot be liable to defensive harm unless the imposition of that harm is necessary to serve a sufficiently just cause, such as defending an innocent person from a threat.¹⁷ Thus, Sam's only reasonable option for surviving in the *present moment* is to kill Meg before she shoots him. Fourth, the threat does not have a just reason for threatening the defender's life. In the absence of a sufficient moral justification on the part of the *unjust threat*, the defender has the right to protect his own life. A threat is required to morally justify her threatening status because there exists a “wrong”: that is, the defender's rights are violated. In this sense, the defender has a moral claim against the threat.

A second approach to justifying killing in self-defense is what I describe as the *culpable threat account*. This differs from the unjust threat account in that it requires that the threat be *morally culpable* for his attack on the defender. Like Thomson's unjust threat account, the culpable threat account for justified killing in self-defense holds that a defender is entitled to kill in self-defense when his life is threatened and he will be killed unless he kills his attacker first. Unlike Thomson's account, however, the culpable threat account insists that moral fault must be attributable to the attacker to justify self-defense. According to Miller, it is not sufficient that there is a deadly threat that can only be removed by killing the person who constitutes the deadly threat. By moral fault, he means that the attacker intends to kill the defender and he is responsible for having this intention to kill the defender.¹⁸ If it can be established that the attacker is morally culpable for his deadly threat to the defender, then this account concludes there is the required “moral asymmetry” between the attacker and the defender.

To illustrate, consider the difference between Thomson's *Villainous Aggressor* and *Innocent Threat* cases. The truck driver in *Villainous Aggressor* is intentionally and unjustly trying to kill the defender. So both accounts agree that the defender is morally justified in destroying the truck and killing the driver in self-defense. But the two accounts disagree about the falling fat man. Although the deadly threat posed by the falling fat man in *Innocent Threat* is unjust, he is not intending to harm the defender. This means the falling fat man is not morally at fault and so the culpable threat account says that killing him in self-defense is not justified. In contrast, on the unjust threat account, it is enough that the falling fat man lacks a justification for posing a deadly threat in order for the defender to be justified in killing him. In short, Miller's culpable threat account agrees that Sam is justified in killing Meg in self-defense when the same four basic conditions are met. But then it adds a fifth condition by requiring that Meg intended to attack Sam and she is responsible for that intention (*culpability condition*).

But moral culpability should *not* be included as a condition because justified killing in self-defense is not based on desert of punishment. Thomson, for instance, rejects two possible reasons for why it might be thought justified for the defender to blow up the truck in *Villainous Aggressor* but not in *Innocent Aggressor*. First, the villainous aggressor might be judged to be less "worthy" than the fault-free driver. Thomson argues against this conclusion on the basis that the fault-free driver might be a much less morally worthy person, all things considered.¹⁹ Second, it might be concluded that the villainous aggressor deserves to be punished whereas the fault-free driver does not. In this case, Thomson makes the point that it is not up to the private person defending herself to mete out punishment in such situations.²⁰ Furthermore, Leverick points out that criticism of the rights-based approach to killing in self-defense has tended to focus on discomfort with the idea that the right not to be killed can be temporarily forfeited. This is, she suggests, especially the case with an innocent aggressor who forfeits the right through no fault of their own.²¹ But when rights forfeiture is not linked to fault, because it is not based on the notion of culpability, then the right not to be killed is only forfeited by virtue of becoming an immediate threat to the life of another. So the unjust threat approach avoids these criticisms because it is *not* saying the threat *deserves* to be killed or is being punished for the threat he poses to the defender.

A third type of rights-based account for justifying killing in self-defense that I'd like to examine here is what I call the *responsible threat account*. McMahan says that in order to overcome or defeat the presumption against intentional killing the threat must have done something morally decisive enough to make him liable to be killed.²² He claims that the threat becomes liable when he "voluntarily engaged in a risk imposing activity and is responsible for the consequences when the risks he imposed eventuate in harms."²³ And when it is established that the threat is to a sufficient degree responsible for posing a threat of unjust harm to the defender then the threat has made himself liable to defensive harm. McMahan points out that the relevant difference between the threat and the defender might be of comparatively slight moral significance. So he suggests that in cases where the costs cannot be divided (i.e., either one or the other will be killed) then he argues it is fair that the threat who has voluntarily engaged in a risk-imposing activity should

suffer death rather than impose it on the defender.²⁴ But McMahan's responsible threat account does not require the threat to be morally at fault or intending to harm the defender for killing in self-defense to be justified. In this way it differs from the culpable threat approach. Returning again to Thomson's hypothetical cases above, the truck driver in *Villainous Aggressor* is intentionally and unjustly threatening to kill the defender. So all three accounts agree that the defender is morally justified in destroying the truck and killing the driver in self-defense, which seems to be the right result. In contrast, the deadly threat posed by the falling fat man in *Innocent Threat* is unjust so the unjust threat account says that the defender is morally justified in killing him in self-defense. But because the falling fat man is neither morally at fault nor responsible for the threat he poses, both the culpable threat account and responsible threat account say that killing in self-defense is not justified. Turning to *Innocent Aggressor*, however, we find that the culpable threat account does *not* justify killing in self-defense because the truck driver is not morally at fault for the deadly threat he poses to the defender. The unjust threat account concludes that killing in self-defense is justified because the deadly threat posed by the truck driver is not justified. The responsible threat account agrees that killing in self-defense is justified in this case but not because the threat is unjust or because the truck driver is morally at fault. Rather, it is because the truck driver is engaging in the risk-imposing activity of truck driving and he is responsible for ensuring the truck does not threaten any innocent lives. In short, McMahan's responsible threat approach also agrees with the four basic conditions but again adds a fifth condition requiring that Meg is responsible for the risk-imposing activity that threatens Sam's life.

McMahan's responsible threat account agrees with the unjust threat account that liability to defensive harm is not based on desert of punishment. But it then disagrees on the question of where the onus for moral justification lies. The unjust threat account of killing puts the "burden of proof" on the threat to justify the danger he poses to the defender. It says that the threat must justify the deadly danger she poses to the defender in order to *not* be liable to defensive harm. The defender, for his part, must prove that his own intentions are defensive when confronted with the reality of an unjust immediate deadly threat. In contrast, the responsible threat and culpable threat accounts demand that the defender proves some additional quality in the threat to establish the necessary moral asymmetry between the two parties. The "burden of proof" then lies with the defender to justify his actions by demonstrating the threat has this asymmetric quality. For the culpable threat account, this means the defender must prove that the threat was morally at fault because he intended the attack. Whereas the responsible threat account demands the defender prove that the threat was sufficiently responsible for the risk-imposing activity that threatens the defender's life.

3. A FORCED CHOICE BETWEEN LIVES

So far, I have highlighted a key point of disagreement between the unjust threat, responsible threat and culpable threat accounts for justifying killing in self-defense. Now I argue that the conditions added by the culpable threat account (*culpability*

condition) and responsible threat account (*liability condition*) are unfair to the defender because they put the “burden of proof” on him to prove something about the threat’s moral state. I then go on to argue that these additional conditions are also largely unnecessary. The unjust threat account for justifying killing in self-defense is sufficient where one faces a *strict* forced choice between lives. According to Montague, a forced choice between lives situation is one where Person B will certainly kill Person A unless Person B is killed first. In other words, he suggests that if the threat is not prevented from doing so, he will in fact kill the defender.²⁵ For example, let’s imagine a hypothetical case where Hans attempts to *intentionally* push John off the edge of a tall building. If Hans is successful then the fall onto the pavement below will certainly kill John. In order to prevent Hans pushing him off the building edge, John grapples with Hans. But Hans is physically stronger and will inevitably win out. The only choice that John has, if he wants to survive, is to twist his body in a way that forces Hans to go off the building edge instead. So either Hans or John will certainly die. In this case, all three accounts agree that John is morally justified in killing Hans in self-defense. Hans is unjustly threatening John’s life, his attack is a risk-imposing activity and he intends to harm John.

But now let us consider a hypothetical case where the threat is innocent of wrongdoing. Imagine that Martin and Roger are window cleaners working high-up on the outside of a building. Martin stumbles and, as he loses his balance, he instinctively grabs at Roger to steady himself. Unfortunately, this action causes Roger to also lose his balance and consequently Roger is faced with two choices. He can do nothing and allow himself to fall from the cleaner’s platform and the fall will kill him. Or he can firmly push Martin away, which will cause Martin to fall and be killed instead. The unjust threat account says that Roger is morally justified in pushing Martin away because he is warding-off a threat to his life. In contrast, both the culpable threat and responsible threat accounts hold that it is *not* morally justifiable to kill either a non-responsible threat or an innocent threat in self-defense because the threat, in this case, is not liable for his threatening action.²⁶ McMahan argues that if the defender’s life is threatened by a non-liable threat-so that he must choose between intentionally killing the threat and allowing himself to be killed by him-the presumption opposes killing in self-defense.²⁷ It is integral to the culpable threat and responsible threat accounts of justified killing in self-defense that there is a sufficient moral difference between the threat and the defender. That is, “the threat must have done something morally decisive enough to make them liable to be killed.”²⁸ So, according to both the culpable threat and responsible threat accounts, Martin should allow himself to fall from the cleaner’s platform rather than push Roger away.

Why should Roger’s life, however, be preferred to Martin’s life in this case? In a situation where there is a forced choice between two innocent lives, it is unfair that the defender is obliged to sacrifice his life in this way. If we accept the conclusion that Roger is not justified in defending his life against the threat from Martin, then this means that choosing to defend himself makes Roger liable to be killed instead. That is, if Roger attempts to intentionally push Martin away, then Martin is morally justified in killing Roger. But this conclusion cannot be correct.

It is absurd that the defender would make himself liable to being killed by the threat because he attempted to defend himself from the threat. Roger's intention, in this case, is to protect himself from an unjust immediate deadly threat rather than to kill his work colleague.

It is also unfair that the burden of proof is on the defender to justify his actions rather than on the threat. As I explained above, both the culpable threat and responsible threat accounts oblige the defender to prove the threat has an additional quality other than the fact that he is posing an unjust immediate deadly threat. The culpable threat account demands proof that it is an intentional attack and the responsible threat account demands proof that the threat was sufficiently responsible for the risk-imposing activity that threatens the defender's life. But why should the burden of proof be on the defender to prove the threat's sufficient intentionality or responsibility for a risk-imposing activity and not on the threat to justify his deadly status? If we take seriously the notion that the defender has a right not to be killed, then it is reasonable to conclude that a person posing an immediate deadly threat to the defender should be required to justify himself. For the defender's part, he should only have to prove that his own intentions are defensive when confronted with the reality of an immediate deadly threat.

It has been argued that cases where the threat is incapable of exercising moral agency, so that he is merely an object, prove that the unjust threat account is wrong. Noam Zohar, for example, suggests that a falling body is the equivalent of a falling piano. A piano cannot commit a rights violation, he suggests, because it is an object and not a moral agent.²⁹ Uwe Steinhoff similarly makes the point that an obligation to justify being a threat would *not* make sense because one cannot be obliged to do things which are beyond one's control.³⁰ So I can have no right that the falling body not kill me and, Zohar argues, if the person, as moral agent, is not about to violate my right not to be killed, then it is misleading to say that he is "about to kill me." Therefore, he concludes that self-defense cannot serve as the grounds for permitting the deflection, unless we are prepared to broaden the notion of self-defense to permit any destruction of another to buy one's own life.³¹ This argument is incorrect, however, because if a mere object, such as a piano, is threatening my life then moral justification using rights forfeiture is not required (or at least not appropriate as a moral justification). But rights forfeiture is the appropriate approach for justifying killing a human being in self-defense. A human being is not merely an object in the same way as a piano. Human beings are rights holders. Perhaps, if the person was already dead then they would be an object in the same way as the piano. But the person acting in self-defense does not need to justify destroying a dead body in the same way as a live human being. The relevant moral difference is that a live person is unjustly posing a deadly threat to the defender. According to Victor Tadros, there is sense in saying that the threat's duty to not harm others persists even when he is unable to perform it.³² How can it matter whether a person's body is the threat if their body is not under the control of their own agency? Tadros answers this question by suggesting that a person has a special responsibility to ensure that his body is not the source of a threat.³³ In general, Tadros argues, a person has much more control over the threats that his own body imposes than others do. This means that if we

each have to bear a greater cost to avert the threat that our own bodies impose on others than we will nevertheless normally have control over which costs we have to bear, as well as how and when they are imposed.³⁴ So he concludes that it is plausible to justify killing the innocent threat in self-defense on the grounds that everybody has a primary responsibility for ensuring that his body does not impose deadly threats on others.³⁵

In addition, I argue that it is misleading to use *strict* forced choice between lives scenarios to ground moral justifications for killing in self-defense. This is because a case of strict forced choice between lives is atypical. That is, it is less likely to occur than a case where the defender has some leeway to accept risk and cost. A strict forced choice between lives case is where the defender *cannot*—by accepting some risk or cost—increase the likelihood of the threat surviving and yet can do enough to preserve his own life if he kills the threat first. For example, referring back to the case of Martin and Roger in the previous section, it is *certain*: 1) Roger will be killed (by Martin) if Roger does nothing; and 2) Roger can preserve his own life *if and only if* he kills Martin first. But these types of cases —where the threat will certainly kill the defender unless he is killed first—are *atypical*.³⁶ It requires an unusual set of circumstances to be certain (or almost certain) that the defender will be killed (by the threat) if the defender does nothing and the defender can preserve his own life *if and only if* he kills the threat first. More often than not, the defender will have some leeway to increase the likelihood of the threat surviving if he chooses to either accept some risk of being killed or some cost to himself. For example, Roger has no intention of harming Martin, so rather than pushing Martin to his death perhaps Roger can see a nearby ledge just below them to which he can jump instead. This choice might be riskier for Roger (he might misjudge his jump and fall to his death) or it might be costly if he lands awkwardly on the ledge and breaks his leg. But the cost and/or risk to Roger might be acceptable if it means both parties survive.

In the above hypothetical case, *cost* is the loss that either party suffers in a threatening situation where persons are harmed and/or property destroyed. Although my main concern so far has been decision-making about killing and the loss of life, there are many lesser (albeit still serious) degrees of loss. McMahan acknowledges that it might be reasonable to demand that the defender share the costs with the threat, if possible, when threat is not culpable for the threat of unjust harm he poses.³⁷ He suggests that if the defender can defend his life in a way that inflicts non-lethal harm on the threat and allows her action to inflict a non-lethal injury on him, then he might be obliged to suffer this non-lethal harm rather than kill the threat. McMahan agrees that killing the threat is wrong when such an alternative is available, particularly when the harm suffered by the defender would not be severe.³⁸ For example, let us imagine a hypothetical case where James is driving his car along the highway when the driver of a truck on the other side of the road heading in the opposite direction—Natasha—loses control of her vehicle and swerves towards James. If James does nothing (or attempts to brake), then the truck will collide with him head-on and likely kill him. But, as it happens, James's car is armed with an anti-tank missile so he has the option to defend himself by destroying the truck and killing Natasha. Although James (in

this case) cannot brake in time to avoid hitting the other vehicle, perhaps he does have the option to swerve and miss Natasha. But this means that he will drive off the side of the road and hit a tree. If James chooses this third option, the cost to him is that his car is wrecked.

Risk, on the other hand, is the probability of harm being done to one (or both) of the parties and the likely seriousness of that harm. Sven Hansson suggests that definitions of the word “risk” generally have two major characteristics in common: it denotes something undesirable and indicates lack of knowledge.³⁹ He also believes that the definition of risk as “expected utility” is better than the definition of risk as “probability” because it includes the severity of the negative outcome, which is a major factor that influences our assessments of risk.⁴⁰ So risk is not merely the likelihood of a bad outcome, it also takes into account severity. A defender might have the choice to increase or lessen the risk to themselves, or to others, depending on their actions. For example, in the case above, the cost of a wrecked car is not James’ only consideration in deciding whether to swerve or not. If James chooses to swerve, then he is more likely to be harmed by hitting a tree than if he chooses not to swerve and uses his anti-tank missile to destroy the truck instead. The cost to James might be a wrecked car but there is also a risk that he will be seriously injured, or even killed, by hitting a tree at high speed.

If it is true that strict forced choice between lives cases are atypical, then it is a misleading place to start. We should not build our understanding of morally justified killing in self-defense on one type of case if another type of case is much more likely to occur in fact. There have been criticisms of the use of hypothetical cases for ethical analysis along these lines. Michael Davis, for example, has cautioned against the excessive use of imaginary cases by philosophers. He says that the more the case departs from this world the greater the chance it will prove to be flawed.⁴¹ Rather than conclude that imaginary cases should be ruled out of ethics, however, Davis suggests that certain routines should be followed when they are used, such as choosing the more realistic version of a case where one has a choice between it and a less realistic one.⁴² Likewise, I have no in-principle objection to using hypothetical cases for the purposes of illustrating a point (as I have demonstrated on a number of occasions in my own work). But one should be skeptical about the more unlikely philosophical cases. It is implausible to think, for instance, that the average person will normally have access to an anti-tank missile capable of stopping a truck dead in its tracks.⁴³ And if we take away his truck destroying capability, then the most likely options for James in the hypothetical case above is to either collide with the truck or swerve and take his chances with the trees. In this case, he has no capability to stop the truck or harm Natasha. But if it turns out that James does have the capability to destroy an out-of-control truck speeding towards him (which is no easy feat), then it seems just as likely that this capability will also give him additional reasonable options to stop the truck that are also less likely to kill the driver. For example, James might aim his missile at one set of wheels, rather than the centre of the truck, so that the speeding vehicle veers away and Natasha has some chance of survival.

In addition, the likelihood of the right set of circumstances necessary for a strict forced choice between lives scenario is lessened even further in cases where the

threat is not intending to harm the defender. The culpable threat who is intending to harm the defender can renew his attack after his initial threat has been foiled and might persevere in his efforts to harm the defender over time. In contrast, the threat that is not intending to kill the defender will *not* persevere in posing a threat in the same way. Consequently, it is unlikely (or at least less likely) that the defender will confront a case where: 1) it is *certain* that the defender will die; 2) unless he kills the threat first; 3) the defender has no other reasonable options that increase the threat's chance of surviving; and 4) the threat is not intending to harm the defender. Instead, the defender is more likely to confront a situation where he either has no capability to stop the threat by killing him or he has a number of options that involve varying levels of cost and risk to himself. If this is true, then the moral justification for killing in self-defense should normally include calculations of risk and cost to both the threat and the defender.

4. A VARIABLE CULPABILITY CONDITION

If there is leeway for the defender to take on cost and risk, then, as I now demonstrate, moral culpability plays a key role in decision-making. Thomson's unjust threat account holds that moral culpability in an attacker – what Thomson refers to as the culpable aggressor – is irrelevant for morally justifying killing in self-defense in a strict forced choice between lives case. She says that “if the aggressor will (certainly) take your life unless you kill him, then his being or not being at fault for his aggression is irrelevant to the question whether you may kill him.”⁴⁴ For Thomson, establishing the threat's culpability is *not* a necessary condition for morally justifying cases of killing in self-defense. As I argued above, this is correct for a strict forced choice between lives scenario. But Thomson acknowledges that there might be room for argument about the role of culpability between “the extremes of very grave bodily harm on the one hand, and loss of wallet or hat on the other hand.”⁴⁵ She deliberately leaves aside the question of whether the threat's culpability makes a difference when it comes to the defender's obligation to share costs. Thomson also leaves open what should be said about cases in which it is not certain that the threat will cause the defender harm.⁴⁶ She argues that “fault is also irrelevant when the aggressor would otherwise blind you, or cut off your legs: the aggressor's fault or lack of fault has no bearing on whether you may kill the aggressor to defend your eyes or legs.”⁴⁷ In short, Thomson insists that the culpability of the threat simply plays no role at all in the defender's moral justification for using force.

Thomson's approach here is a mistake, however. In some situations, culpability clearly plays some sort of role in justifying decisions to use lethal force. Robert Fullinwider, for example, describes a case where mobsters kidnap Smith's children and threaten to kill them unless he kills Jones. Driven by the threat, Smith seeks out Jones to shoot him. In the place that Smith finds Jones, the unarmed mobsters are parked across the street to make sure he goes through with the mob hit.⁴⁸ Fullinwider's point is that, after killing Smith, Jones could not then turn his gun on the mobsters (perhaps because he realised they were unarmed). Despite the mobsters' culpability in the attack on Jones, it was only Smith who

was the agent of immediate threat to Jones. The mobsters were not posing a direct and immediate danger and so, he argues, should not be killed in self-defense.⁴⁹ Lawrence Alexander agrees that after killing Smith, Jones may not invoke self-defense to then turn and kill the mobsters. At most, he suggests, the mobsters may be punished for their guilt in instigating the murder.⁵⁰ But, he asks, “may Jones invoke the Principle of Self-Defense to kill the mobsters instead of Smith if by doing so he will cause Smith to relent?” If the mobsters had a gun trained on Smith and had ordered him to kill Jones, and he were about to comply, then Alexander argues that Jones not only could, *but should*, kill the mobsters rather than Smith.⁵¹ This is because the mobsters and Smith are both necessary causes of the danger to Jones (and killing either the mobsters or Smith removes the danger). If true, then Jones should kill the ones who are morally at fault for the attack (the mobsters) rather than killing the innocent attacker (Smith).⁵² The point here is that culpability matters.

Another limitation of Thomson’s unjust threat account is her argument that intention is irrelevant for morally justifying killing in self-defense. Thomson disagrees with the distinction, made in the literature on the doctrine of double effect, between foreseen and intended effects arguing that “if fault is irrelevant to permissibility, then so also is intention.”⁵³ She then describes a hypothetical case to support her contention. In this case, Thomson explains that Alfred’s wife is dying and he wishes to hasten her death by poisoning her. But he doesn’t know that the poison is actually the only existing cure for her ailment. Thomson suggests that it is permissible for Alfred to give his wife the poison because “how could his having a bad intention make it impermissible for him to do what she needs for life.”⁵⁴ But this is incorrect because if it turns out that Alfred’s true intention is revealed to the authorities (perhaps he told a third-party about his plan to kill his wife and they reported him to the police), then he would be guilty of attempted murder. The fact that Alfred inadvertently heals his wife, rather than killing her, would not be relevant in our condemnation of his action in attempting to poison his wife.⁵⁵ In all cases of genuine self-defense, suggests Uniacke, the agent’s aim is to stop the threat. Strictly-speaking, the death of the threat is not required to achieve this intended goal.⁵⁶ In justified self-defense, the defender’s primary intention must be to repel the threat to his life rather than kill the threat. If true, this means the defender’s intention is a necessary feature of the defense condition.

So what does this mean for the unjust threat account of justified killing in self-defense? It means that the unjust threat account should be modified to incorporate the role of the threat’s moral culpability (i.e., the threat’s moral fault) in judgments about justified killing in self-defense. It can do this by acknowledging that there are degrees of liability to defensive harm. According to Bradley Strawser, this suggests that the threat’s liability to defensive harm varies according the degree to which she is culpable for the threat she poses.⁵⁷ This is consistent with McMahan’s way of approaching the real-life complexity of killing in self-defense and the role of culpability. He suggests that the harmfulness of the defensive action to which the partially-excused threat is liable varies with the degree of her moral culpability.⁵⁸

It also means that two factors that Thomson puts aside—cost and risk—play an important role in how morally justified killing in self-defense works in practice. The defender's obligation to share the cost and risk of harm will vary depending on whether the threat is non-culpable, partially-culpable or fully-culpable. This is based on the notion that the defender's obligations to take on risk and cost are contingent on the threat's culpability. In cases where the threat is non-culpable, the defender is strongly obliged to share the cost and risk. If, in the process of defending himself, the defender has an opportunity to preserve the life of the innocent threat then he should do so up to the point where the risk or cost is about the same for both parties. Neither party is morally blameworthy for the situation, so fairness suggests that both parties should share the misfortune equally (if that is possible). In the case of the *Innocent Threat*, the defender should choose to accept a severe injury if it means the innocent fat man will survive. The defender might also judge that he has some chance of surviving if the fat man falls on him, whereas the fat man will have no chance of survival if he is deflected away. In this case, the defender should accept the risk of being killed.

In cases where the threat is partially-culpable, however, the defender has a lessened obligation to share some of the cost and risk, including the potential for serious injury. That is, the defender should share the cost and risk when the threat is non-culpable or partially-culpable but the degree varies. This is the *variable culpability condition*. This says that the defender's obligation to share the cost and risk required for the threat to survive varies according to the culpability of the threat. In the case of the *Innocent Aggressor*, the defender is justified in destroying the truck, but he has an obligation to accept a moderate amount of cost and/or risk to preserve the life of the truck driver. So the defender might be obliged to accept a broken arm but not a severe injury that leads to a permanent disability.

Finally, the defender is only obliged to accept a negligible amount of cost and/or risk to preserve the life of the threat who is fully-culpable. The obligation exists but it is of a much weaker variety. For example, a thief might snatch a woman's purse and run off. This would cause her an inconvenient loss (e.g., she loses her purse) and there might be a slight risk of harm (e.g., perhaps she falls over and grazes her knee). It would not be justified to kill a thief who poses this type of threat, however.⁵⁹ In cases where the danger from the attacker is negligible, the defender should accept the minor loss and/or very slight risk of harm rather than kill him. But this obligation to accept some risk or cost from a culpable attacker is much weaker than the obligation that applies to the threat who is either non-culpable or partially-culpable. Hence, in the case of the *Villainous Aggressor*, the defender is justified in destroying the truck with an obligation to accept only a negligible amount of cost and/or risk to preserve the life of the truck driver. In short, the defender should apply a *variable culpability condition* in cases where the threat is non-culpable or partially-culpable. This says that the defender's obligation to share the cost and risk required for the threat to survive decreases as the threat's culpability increases. This means that Meg's culpability for the threat she poses to Sam weakens Sam's obligation to share the cost and risk required for both parties to survive (*variable culpability condition*). In other words, I am advocating a different position to the culpability or liability-based accounts.

5. CONCLUSION

The best explanation for morally justifying killing in self-defense from the rights-based perspective is a modified unjust threat account. This says that killing in self-defense is morally justified by the necessity of repelling or warding off an unjust immediate deadly threat. In addition, the general thrust of McMahan's argument is correct: the obligation to be restrained in the use of force, and accept higher levels of risk and cost, has an inversely proportional relationship to the moral culpability of the attacker. This means that in a choice between the lives of two innocent persons there is an obligation on the defender to first seek to share the cost and risk between himself and the threat who is non-culpable, if it means that both are more likely to survive. Failing that, the unjust threat's moral obligation to *not* pose a deadly threat to the defender creates the moral asymmetry necessary to justify the defender killing in self-defense. In cases where the unjust threat is culpable (or partially culpable), the obligation to share cost and risk diminishes accordingly. If this is correct, then we can conclude that the defender (Sam) is justified in killing the unjust threat (Meg) in self-defense when all the following conditions are met: 1) Meg is an immediate deadly threat to Sam (*immediate threat condition*); 2) Sam's primary intention is to preserve his own life, which can be achieved by killing Meg (*defensive intent condition*); 3) Sam's only reasonable option for preventing being killed is to kill Meg (*necessity condition*); 4) Meg does not have a sufficiently just reason for threatening Sam's life (*unjustified condition*). If there is leeway, however, for the defender to take on costs and/or risks then 5) Sam is obliged to share the cost and/or risk equally with Meg for both parties to survive (*shared costs and/or risks condition*) but 6) Meg's culpability for the threat she poses to Sam weakens Sam's obligation (*variable culpability condition*).

Competing interests: The author declares none.

NOTES

1. Although rights-based approaches are not the only form of justificatory reasoning for killing in self-defense, they are the ones I'm choosing to concern myself with here. See for example: Suzanne Uniacke, *Permissible Killing: The Self-Defence Justification of Homicide*, "Cambridge Studies in Philosophy and Law" (Cambridge, UK: Cambridge University Press, 1994); David Rodin, *War and Self-Defense* (New York: Oxford University Press, 2003); Jonathan Quong, "Killing in Self-Defense," *Ethics* 119, no. 3 (2009); Fiona Leverick, *Killing in Self-Defence* (Oxford: Oxford University press, 2006); Helen Frowe, *Defensive Killing* (Oxford: Oxford University Press, 2014); Uwe Steinhoff, *Self-Defense, Necessity, and Punishment: A Philosophical Analysis* (Taylor & Francis, 2019).

2. Judith J. Thomson, "Self-Defense," *Philosophy & Public Affairs* 20, no. 4 (1991).
3. Seumas Miller, "Killing in Self-Defense," *Public Affairs Quarterly* 7, no. 4 (1993).
4. Jeff McMahan, *Killing in War* (Oxford: Oxford University Press, 2009).
5. Rodin, *War and Self-Defense*, 2.
6. Leverick, *Killing in Self-Defence*, 44.

7. Uniacke, *Permissible Killing: The Self-Defence Justification of Homicide*, 196.
8. Miller, *Killing in Self-Defense*, 327.
9. Uniacke, *Permissible Killing: The Self-Defence Justification of Homicide*, 184.
10. Thomson, *Self-Defense*, 301.
11. Ibid. 283.
12. Ibid. 284.
13. Ibid.
14. Ibid. 287.
15. Ibid.
16. Ibid. 288.
17. Joanna Firth and Jonathan Quong, "Necessity, Moral Liability, and Defensive Harm," *Law and Philosophy* (2012): 674.
18. Miller, *Killing in Self-Defense*, 331.
19. Thomson, *Self-Defense*, 285.
20. Ibid.
21. Fiona Leverick, "Defending Self-Defence," *Oxford Journal of Legal Studies* 27, no. 3 (2007): 572.
22. McMahan, *Killing in War*, 169.
23. "The Basis of Moral Liability to Defensive Killing," *Philosophical Issues* 15, no. 1 (2005): 394.
24. Ibid.
25. Phillip Montague, "Self-Defense and Choosing between Lives," *Philosophical Studies* 40, no. 2 (1981): 210.
26. Working as a window cleaner high up on the side of building is a risky activity for the cleaners but **not** is not risk-imposing in the sense that driving a car imposes significant risks to pedestrians, cyclists or the occupants of other cars.
27. McMahan, *Killing in War*, 169.
28. Ibid.
29. Noam J Zohar, "Collective War and Individualistic Ethics: Against the Conscription of 'Self-Defense,'" *Political Theory* 21, no. 4 (1993): 608.
30. Uwe Steinhoff, *On the Ethics of War and Terrorism*, (Oxford: Oxford University Press, 2007), 82.
31. Zohar, "Collective War and Individualistic Ethics: Against the Conscription of "Self-Defense"," 608.
32. Victor Tadros, *The Ends of Harm: The Moral Foundations of Criminal Law*, (Oxford: Oxford University Press, 2011), 255.
33. Ibid., 252–53.
34. Ibid., 253.
35. Ibid., 252–53.
36. I acknowledge that the hypotheticals employed by philosophers are oftentimes deliberately atypical: ethics is done at the margins so to speak. But the focus on atypical cases in philosophy **is** problematic.
37. McMahan, *The Basis of Moral Liability to Defensive Killing*, 394.
38. Ibid.

39. S.O. Hansson, *The Ethics of Risk: Ethical Analysis in an Uncertain World* (Palgrave Macmillan, 2013), 4.
40. *Ibid.* 7.
41. Michael Davis, "Imaginary Cases in Ethics," *International Journal of Applied Philosophy* 26, no. 1 (2012): 14.
42. *Ibid.* 6.
43. McMahan, *The Basis of Moral Liability to Defensive Killing*, 393.
44. Thomson, *Self-Defense*, 286.
45. *Ibid.*
46. *Ibid.*
47. *Ibid.*, 285.
48. Robert K. Fullinwider, "War and Innocence," *Ibid.* 5, no. 1 (1975): 92.
49. *Ibid.*, 93.
50. LA Alexander, "Self-Defense and the Killing of Noncombatants: A Reply to Fullinwider," *Ibid.* no. 4 (1976): 409.
51. *Ibid.*
52. *Ibid.*, 410.
53. Judith J. Thomson, "Self-Defense," *Ibid.* 20 (1991): 295.
54. *Ibid.*, 294.
55. Thomson also does not make clear Alfred's ultimate motivation for attempting to kill his wife. Is he maliciously trying to get rid of her or does he love her and believes he is mercifully ending her inevitable suffering? If it is the latter, then we might say that while Alfred's immediate intention was to kill his wife, given all the facts he would choose to save his wife. In this case, his immediate intention to kill means his action is still morally impermissible but his ultimate motivation makes for a potentially exculpatory excuse.
56. Uniacke, *Permissible Killing: The Self-Defence Justification of Homicide*, 108.
57. Bradley Jay Strawser, "Walking the Tightrope of Just War," *Analysis* 71, no. 3 (2011): 540.
58. According to McMahan, a partially-excused threat is a person who unjustifiably poses a threat of wrongful harm to others but whose actions are excused to some extent without being fully justified. *Killing in War*, 159–61.
59. Of course, the defender still has a proportionate right to defend herself and her property. It would be permissible, for example, for the defender (or third party) to yell at the thief, push him away or tackle him to the ground in order to prevent the theft.

REFERENCES

- Alexander, LA. "Self-Defense and the Killing of Noncombatants: A Reply to Fullinwider." *Philosophy & Public Affairs* 5, no. 4 (1976): 408–15.
- Davis, Michael. "Imaginary Cases in Ethics." *International Journal of Applied Philosophy* 26, no. 1 (2012): 1–17. <https://doi.org/10.5840/ijap20122611>
- Firth, Joanna, and Jonathan Quong. "Necessity, Moral Liability, and Defensive Harm." *Law and Philosophy* (2012): 1–29. <https://doi.org/10.1007/s10982-012-9135-4>

- Frowe, Helen. *Defensive Killin*, Oxford: Oxford University Press, 2014.
<https://doi.org/10.1093/acprof:oso/9780199609857.001.0001>
- Fullinwider, Robert K. "War and Innocence," *Philosophy & Public Affairs* 5, no. 1 (1975): 90–97. <https://doi.org/10.2307/j.ctv24rgbr1.7>
- Hansson, S.O. *The Ethics of Risk: Ethical Analysis in an Uncertain World*, Palgrave Macmillan, 2013.
- Leverick, Fiona. "Defending Self-Defence," *Oxford Journal of Legal Studies* 27, no. 3 (2007): 563. <https://doi.org/10.1093/ojls/gqm012>
- Leverick, Fiona. *Killing in Self-Defence*,. Oxford: Oxford University press, 2006.
<https://doi.org/10.1093/acprof:oso/9780199283460.001.0001>
- McMahan, Jeff. "The Basis of Moral Liability to Defensive Killing," *Philosophical Issues* 15, no. 1 (2005): 386–405.
- McMahan, Jeff. *Killing in War*, Oxford: Oxford University Press, 2009.
<https://doi.org/10.1093/acprof:oso/9780199548668.001.0001>
- Miller, Seumas. "Killing in Self-Defense." *Public Affairs Quarterly* 7, no. 4 (1993): 325–39.
- Montague, Phillip. "Self-Defense and Choosing between Lives," *Philosophical Studies* 40, no. 2 (1981): 207–19. <https://doi.org/10.1007/BF00353792>
- Quong, Jonathan. "Killing in Self-Defense," *Ethics* 119, no. 3 (2009): 507–37.
<https://doi.org/10.1086/597595>
- Rodin, David. *War and Self-Defense*, New York: Oxford University Press, 2003.
<https://doi.org/10.1093/0199257744.001.0001>
- Steinhoff, Uwe. *On the Ethics of War and Terrorism*, Oxford: Oxford University Press, 2007.
<https://doi.org/10.1093/acprof:oso/9780199217373.001.0001>
- Steinhoff, Uwe. *Self-Defense, Necessity, and Punishment: A Philosophical Analysis*, Taylor & Francis, 2019. <https://doi.org/10.4324/9780367814441>
- Strawser, Bradley Jay. "Walking the Tightrope of Just War," *Analysis* 71, no. 3 (July 1, 2011 2011): 533–44. <https://doi.org/10.1093/analys/anr070>
- Tadros, Victor. *The Ends of Harm: The Moral Foundations of Criminal Law*, Oxford: Oxford University Press, 2011. <https://doi.org/10.1093/acprof:oso/9780199554423.001.0001>
- Thomson, Judith J. "Self-Defense." *Philosophy & Public Affairs* 20, no. 4 (1991): 283–310.
- Uniacke, Suzanne. *Permissible Killing: The Self-Defence Justification of Homicide*. *Cambridge Studies in Philosophy and Law*, Cambridge, UK: Cambridge University Press, 1994.
<https://doi.org/10.1017/CBO9780511554483>
- Zohar, Noam J. "Collective War and Individualistic Ethics: Against the Conscriptio of" Self-Defense." *Political Theory* 21, no. 4 (1993): 606–22.
<https://doi.org/10.1177/0090591793021004003>