

Responsibility and Self-Defense: Can We Have It All?

Adam Hosein¹

Published online: 5 October 2016
© Springer Science+Business Media Dordrecht 2016

Abstract The role of responsibility in our common-sense morality of self-defense is complex. According to common-sense morality, one can sometimes use substantial, even deadly, force against people who are only minimally responsible for posing a threat to us. The role of responsibility in self-defense is thus *limited*. However, responsibility is still sometimes *relevant*. It sometime affects how much force you can use against a threatener: less if they are less responsible and more if they are more responsible. Is there a well-motivated theory that can explain both why the role of responsibility is *limited* and why it is sometimes *relevant*? It is hard to see what theory could unify these disparate elements of our common-sense morality, and if one cannot be found then we may simply have to revise some of our pre-theoretic beliefs. But it would be an important advantage of a theory if it could justify those beliefs. I will argue that there is a theory of this kind: surprisingly, the familiar rights theory of self-defense, defended by Judith Thomson, can do so if it is suitably supplemented. Along the way I will survey some alternative theories of self-defense and show why they are not up to the task.

Keywords Self-defense · Responsibility · Beneficence · Killing

Introduction

Suppose that someone is posing a threat to you and you want to know whether you can use defensive force against her. Does it matter how responsible she is for posing that threat? For instance, does it matter whether she is going to harm you intentionally or as a result of an understandable mistake?

✉ Adam Hosein
adam.hosein@colorado.edu

¹ Philosophy Department, University of Colorado, Boulder, UCB 232, Hellems 167, Boulder, CO 80309-0232, USA

The role of responsibility in our common-sense morality of self-defense is complex. In some ways, the responsibility of someone posing a threat to us plays a very limited role. According to common-sense morality, one can sometimes use very substantial, even deadly, force against people who are only minimally responsible, or even non-responsible, for posing a threat. The role of responsibility is thus *limited*. However, responsibility is still sometimes *relevant*. It affects, in some cases, how much force you can use against a threat: less if they are less responsible and more if they are more responsible.

Is there a well-motivated theory that can explain both why the role of responsibility is *limited* and why it is sometimes *relevant*? It is hard to see what theory could unify these disparate elements of our common-sense morality, and if one cannot be found then we may simply have to revise some of our pre-theoretic beliefs. But it would be an important advantage of a theory if it could justify those beliefs. I will argue that there is a theory of this kind: surprisingly, the familiar rights theory of self-defense, in the form defended by Judith Jarvis Thomson (1991), can do so if it is suitably supplemented.¹ Along the way I will survey some alternative theories of self-defense and show why they are not up to the task.

The Rights Theory

Let me begin with a brief reminder of what exactly a theory of self-defense is supposed to explain. Suppose that someone is maliciously attempting to harm me, by jabbing their knife at my chest, say. I may respond by using (necessary and proportionate) force to stop his attack; stabbing him first, say. Why is this? Under normal circumstances this person would have a right against me that I not stab him and this would make it impermissible for me to stab him. So something needs to be said to explain why stabbing him is permissible in this instance.

Perhaps a right not to be harmed can be overridden where the good produced by harming someone is great enough. For instance, perhaps I could punch someone to prevent a thousand deaths. But the present case doesn't seem to be like this. By stabbing him I would only prevent myself from getting stabbed, clearly not a large benefit to override a right of his to not be stabbed. Some philosophers say that it matters here that the good I will secure is a good for myself, rather than others.² I have especially strong reasons to benefit myself, they say, and these reasons weigh against other people's rights. I find this view very implausible: I can steal no more money from you to benefit myself than I can steal from you to benefit a stranger. Similarly, if a threatener has a right not to be stabbed by me, then it is irrelevant that by stabbing him I would benefit myself.³

¹ I draw heavily on Thomson throughout and attempt to show that her view of self-defense is best understood in light of her broader theory of rights.

² For instance, Quong (2009).

³ I realize that 'threatener' may be a neologism, but 'attacker' is too narrow, for reasons that will become clear.

So, it seems that the threatener must lack a right not to be stabbed by me, given that this is a necessary means of stopping him. This is a plausible explanation of why I may harm him even though I would not normally be permitted to harm him. Let us call his lacking a right against me that I not stab him his being ‘liable’ to being stabbed by me (this is a purely stipulative use of the word ‘liable’—if you think its ordinary usage is entirely different that’s fine). And let us call infliction of harm that is a means of stopping the threat posed to me ‘defensive harm’ (there may be other uses of defensive force, such as pre-emptive strikes, that are sometimes permissible, but let us set these aside for now). Normally, people are not liable to being harmed, and so I may not harm them, but the threatener is liable to defensive harm, and so I may stab him.⁴

Most theories of self-defense agree with what I have said so far.⁵ They agree that you may harm a threatener in self-defense, even though harming is normally impermissible, because she is liable to defensive harm.⁶ The remaining question is: what makes her liable? The rights theory of self-defense provides a distinctive answer to this question. It says that a threatener becomes liable to defensive harm when and because she will otherwise wrong someone, or infringe his rights.⁷ The malicious threatener is liable to being harmed in defense by me because she will otherwise infringe my right not to be harmed.

The theory is attractive for its simplicity and for its ability to explain some basic features of the morality of self-defense. For instance, self-defense must be proportionate: the amount of defensive harm I am going to inflict on a threatener must be in proportion to the amount of harm that she would otherwise inflict on me. I may kill you to prevent you from killing me, but I may not kill you to prevent you from cuffing my ear. The rights theory can explain this as follows (it will also be useful later to have the details of this explanation on hand).⁸

Some wrongs are worse than others and thus some rights are more stringent than others. If I steal your pencil I have done you a wrong, but if I steal your prized guitar I have done you a worse wrong. This explains why I would owe you more compensation in the latter case than in the former. It also explains why the amount

⁴ There are in fact a range of ways in which necessary defensive action might involve a response that harms the threatener, including action to block the threat, action to incapacitate the threatener, and so on. For ease, I am going to speak, as McMahan and others do, of liability to harm, but it should be assumed that the harm is being inflicted for one of these specific purposes.

⁵ McMahan (2005), for instance, presents the various major theories of self-defense as in agreement that you can kill in self-defensive when and because attackers are liable to defensive action.

⁶ I am focusing on the claim that it is *permissible* to use defensive force. Kantians might defend a stronger claim that it is *obligatory* to use defensive force, perhaps out of a duty to respect one’s own autonomy. I set this approach aside for present purposes, since it conflicts with the intuitive view that defending yourself is optional and because accepting it would require adopting a broader Kantian theory that I cannot fully consider here.

⁷ I am going to assume that infringing someone’s rights is equivalent to wronging him, though the more basic notion, I think, is that of wronging. And I here use the word ‘infringe’ so that it covers all species of wrongings, including those that are (all things considered) permissible and (all things considered) impermissible. There are various different ways of using the term in the literature, but this usage will make my presentation easier. I believe this formulation of the rights theory is the strongest and I defend it in Hosein (2014).

⁸ This is mostly an elaboration of Thomson (1991, fn. 13).

of benefit to others needed to justify stealing your guitar is much greater than the amount needed to justify stealing your pencil (though, admittedly, on some views no amount of benefit to others can justify either). Similarly, since liability is generated, on the rights view, by threatening to wrong someone, the amount of harm someone is liable to is fixed by how great a wrong they threaten to inflict (or how stringent a right they threaten to infringe): the harm they are liable to is thus proportionate to the gravity of the wrong. And so the more harm someone threatens to inflict the more harm they are liable to, since the more harm you inflict on someone the greater the wrong you do them.

Finally, I think that once we accept that there are such things as rights it is very hard to resist the rights theory. To deny it, one would have to say either, first, that there are situations where A is going to wrong B, by harming him, but B would all the same wrong A by using force to stop this. Surely you do not wrong someone if (within constraints of necessity and proportionality) you use force just to prevent them from wronging you? Or, second, one would have to say, still more implausibly, that there are cases where someone becomes liable to being harmed even though they are neither wronging, nor going to wrong, anyone else (and also have not consented to being harmed). It thus seems to me that the onus is on someone who denies the rights theory to explain why at least one of these apparently implausible claims is in fact true.⁹

The Limited Role of Responsibility

I would now like to turn to the role of responsibility in self-defense cases and what the rights theory says about it. Consider the following examples:¹⁰

Driver

Threatener owns a car, which she maintains carefully, according to all of the manufacturer's instructions. You are pedestrian, walking down the sidewalk at noon. She drives to the grocery store for some supplies and, despite her efforts, the car malfunctions and skids towards you, who will be killed by the impact.

Refill

Threatener has a toy water gun that she generally fills with water. Completely unbeknownst to her, however, someone filled it with acid this morning. She now goes to spray you with it, still thinking that it will merely make you damp. In fact, it will fatally wound you and you know this.

⁹ Thomson does not quite put it like this, but much of the force of her article comes from this challenge.

¹⁰ The **Driver** and **Cell Phone** examples are adapted from those in McMahan (2011, pp. 164–165), while **Falling Person** is a variant on his case in (McMahan 2009).

Cell Phone

A villain has rigged Threatener's cell phone, so that if she calls anyone a bomb strapped to you will detonate. She is completely unaware of this and has no reason to suspect it. She is now dialing someone's number.

Falling Person

You are lying on a deck enjoying the sun with your leg in traction. A sudden and forceful gust of wind blows a person off a ledge above you. She is going to land on you and thereby kill you.

In each of these examples, Threatener poses a lethal threat to you. Suppose further that in each of them you have a weapon with which you can defend yourself from her, that using it will result in her death, and that this is the only way for you to defend yourself. For instance, in **Falling Person**, the fact that your leg is in traction means that there is no way to get out of the way of the falling person in time, but you can save yourself by hoisting your sun umbrella and impaling her on it. And in **Driver** the shotgun you can use to deflate the car's tires so that it stops will also inevitably set its gas tank alight, killing the driver. According to common-sense morality you may fire your weapon in each case.

I would like to now consider the role of responsibility in these examples, so let me first clarify what I mean by 'responsibility'. When I say 'responsible' I do not mean the same thing as 'culpable' (or 'blameworthy' or 'at fault'). As McMahan has shown, responsibility is a broader category than culpability. The **Driver** case illustrates this. Driver has taken all the reasonable precautions that can be expected of someone operating an automobile and so is surely not *culpable* for the ensuing malfunction. All the same, the Driver does seem to be more *responsible* for the situation than, say, the pedestrian. She engaged in the activity of driving full in knowledge that this activity is potentially dangerous to pedestrians, as shown by the number of roadside fatalities each year. So one can be responsible for the threat one poses to others without being culpable for it. The water gun operator in **Refill** is also clearly not culpable, though she is still (minimally) responsible for the threat she poses, given that she is voluntarily doing something which carries some (extremely small) chance of being lethal to others. The falling person is surely neither responsible nor culpable, given her complete lack of agency in bringing about the threat to you. And the same seems to be true of the person making a call in **Cell Phone**, given that the threat she poses is completely unforeseeable.

The examples thus suggest that the role of responsibility (and not just culpability) in the morality of self-defense is quite limited. At least where we are dealing with lethal threats, it does not seem to matter how responsible the threatener is: you can use lethal force against a. minimally responsible threateners, as in **Refill**, and b. non-responsible threateners, as in **Cell Phone** and **Falling Person**.

Can the rights theory explain the limited role of responsibility? It all turns on what rights people have. According to the rights theory, liability is generated when someone is going to infringe a right. Thus, if minimally responsible and non-responsible threats are to be liable to defensive force, then they must be threatening

to infringe a right. For instance, in the cases described above it must be the case that Threatener is going to infringe a right. And indeed the right she is going to infringe must be a *very stringent* one, since killing Threatener is a proportionate response, and proportionality is determined, as I explained above, by how stringent a right someone threatens to infringe. Given that lethal force may be used against her, the natural explanation, on the rights theory, is that she will otherwise infringe the right not to be killed: a clearly very stringent right.

Will she otherwise infringe that right? Is it the case that minimally responsible threateners and non-responsible threateners can infringe people's rights? And is it the case that they can do someone a very *severe* wrong/infringe a very stringent right despite their lack of responsibility? Thomson (1991, pp. 300–301) claimed that we should say 'yes' to all of these questions and several considerations, raised in various parts of her work, support this answer.^{11,12}

A first point in favor of Thomson's view is that it fits with intuitions from outside of the defense context.¹³ To illustrate: if I steal your car, I will infringe your property right in it. What if now someone puts me under duress to steal it, say by threatening to harm me if I do not steal the car? If the duress is sufficiently strong, then it creates an excuse for my stealing the car. And that excuse reduces the degree of responsibility I bear for stealing the car. Yet it seems that I must compensate you for the theft (assuming you cannot recover from the coercer). And the amount of compensation owed does not seem to scale down with my diminished responsibility: significant compensation is still owed by someone whose responsibility is substantially limited by duress.¹⁴ Moreover, tort law takes the same view of such cases: duress is no defense to tort liability (though it is, of course, to criminal liability), and so what compensation I owe (legally speaking) to a victim is not affected by the degree of duress I acted under.

This suggests that in general my degree of responsibility does not determine the wrong that I do you (including whether it is a very significant wrong or a less significant one), since compensation is a way of correcting a wrong done. And if that is right, then it is plausible to think that someone can commit a significant

¹¹ There is, of course, a large literature by now on the relation between rights and intention, responsibility, culpability, and so on. I can not fully survey that literature here, but for our purposes it will be enough to see that Thomson's views on rights are at least one plausible option. I aim to show that with them on hand we can account fully for the role of responsibility in self-defense. This constitutes further evidence for her views about rights. For further discussion about intentions, permissibility, and so on, see, for instance, Scanlon (2008).

¹² Also note that McMahan himself, whose account provides the clearest contrast with Thomson, seems to accept that there is a basic right not to be harmed (a right not tied to the threatener's responsibility). For according to his account you can only defend yourself against a threatener who is responsible for an *unjust* threat. And someone creates an unjust threat of harm just in case they would wrong you (or, equivalently, violate your rights) by inflicting that harm on you.

¹³ Intuitions elaborated on in Thomson (1992).

¹⁴ Though there may be some temptation to demand a little less compensation from someone who has a good excuse for wrongdoing and would be very heavily burdened by having to provide the full amount of compensation demanded of a fully responsible wrongdoer: in so far as this is so, the account I will offer below explains why.

wrong by harming another, even if she bears very little responsibility for inflicting the harm or bears no responsibility for inflicting the harm.

Second, we can give a good explanation for much of the temptation to connect rights infringement closely with responsibility or culpability. The temptation arises from a failure to distinguish clearly between infringing rights and doing something that reflects poorly on one's character (Thomson 1991, pp. 293–296). For instance, suppose that I sell you some tickets to see an artist you enjoy. I have every intention of giving you the tickets and conscientiously attempt to put them in an envelope to send to you. But, making an almost unavoidable mistake, I put in some farmers' market tickets instead which, quite by chance, look almost identical. Since this is an understandable mistake, my blameworthiness and responsibility for failing to send the tickets are diminished. And given my entirely understandable mistake, it is very tempting to say that I have not wronged you, or done you a much less significant wrong than if I had sent the wrong tickets deliberately. But this temptation should be resisted: in fact, my lack of responsibility for this failure is relevant to assessment of my character but not to the fact that I have wronged you. This seems clear once we distinguish carefully between character assessment and assessment of permissibility.

All of this suggests that whether I wrong you, or infringe your rights, by harming you is not determined by how responsible I am for the harm. And if this is right then the rights theory has a strong explanation for why you may defend yourself in **Refill**, **Cell Phone** and **Falling Person**: Threatener is going to infringe your rights, even though she is only minimally responsible or non-responsible. The theory is able to explain the various ways in which common-sense morality gives *limited* significance to responsibility in self-defense cases.

Now, some people will be especially resistant to the claim that there is really a rights infringement in **Falling Person**, and thus resistant to the claim that the rights theory can explain why self-defense is permissible in that case.¹⁵ For our purposes, this does not matter too much, since **Refill** and **Cell Phone** are by themselves quite sufficient to show the advantages of the rights theory (especially over the other options we will consider later). But, while I cannot fully survey the literature here, I will argue that there is in fact a strong *prima facie* case for thinking that there is a rights infringement in **Falling Person**.

According to the skeptic, you can only infringe a right while exercising *agency*: loosely, when you are acting under your own direction.¹⁶ If I raise my arm in order to flag a cab I am exercising agency. If my arm rises up because someone else hoists it by a rope, or because of a spasm, I am not exercising agency.

Here are some replies to the skeptic. First, when we consider examples, there do seem to be moral rights of this kind.¹⁷ For instance, suppose that my car has rolled onto your driveway from mine due to a totally unforeseeable malfunction, and I am

¹⁵ For a very helpful summary of the literature on the rights theory and cases where a threat lacks agency, see Doggett (2011).

¹⁶ As claimed by, for instance, Otsuka (1994).

¹⁷ See Kamm (1992, p. 48) for similar examples involving people's bodies, though I find intuitions are especially clear about the car case described.

currently incapacitated, through no fault of my own, which means that I can do nothing to remove the car.¹⁸ When I am no longer incapacitated, I must take steps to remove my car. And it seems incorrect to say that your rights are only infringed from the moment that I regain my agency: surely the longer the car has been there, the greater the pressure on me to remove it quickly, and this is plausibly because the longer the car has been there the more significantly your rights have been infringed. Also, suppose that I have an assistant who generally acts on my behalf. If, while I am incapacitated, he can take steps to remove the car, he should do so, even if this will result in various costs for me (say because he will have to abandon an important task that I have asked and need him to complete).

These duties—the demands on me when I regain my agency and the demands on my assistant—surely share a common source. And a simple and plausible characterization of that source is a right that you have against me that my car simply not *be* in your driveway. The example suggests, then, that you have a right that is infringed whenever my car is in your driveway, even if it got there without any exercise of my agency and I am currently not able to exercise agency.

The reason why this right exists is, plausibly, that since I own the car I am responsible for it, including its not being somewhere it ought not be, such as your driveway. It is similarly plausible to think that I am responsible for where my body is, given that it is something closely identified with me as a person, and this includes responsibility for my body not being in, as Frances Kamm (1992, p. 47) puts it, the ‘morally inappropriate position’ of causing you harm.¹⁹ There is thus a plausible case for thinking that even in **Falling Person** the victim’s rights are under threat.

Now, it is often more natural to describe situations involving threateners who lack agency as cases where the victim’s rights are *being infringed*, rather than as cases where someone *is infringing* the victim’s rights. In the car case just discussed, for instance, it does seem implausible to say that I am *infringing* your rights by the car’s rolling down the hill, but it does seem plausible, and fits with common-sense notions of property rights and trespass, to say that you have a right that my car not *be* on your property. If so, then the rights theory can be reformulated as the view that A is liable to defensive harm from B when and because B has a right against A, and the defensive harm is a necessary means of preventing that right from being infringed. And this is a plausible formulation of the rights theory, since it retains the basic intuition that people can use force to protect their rights. (Since it is cumbersome, however, I will continue for most of the paper to state the rights theory using the narrower formulation that I began with.)

¹⁸ Someone might say that still you exercised agency in *buying* the car and *parking* it your drive and thus accepted all potential liability risks associated with it. I think if we say that we might as well also say that when you walk outside you assume some risk of falling on someone at the bottom of a well due to freak weather.

¹⁹ I borrow the identification point from Victor Tadros’ (2011, p. 55) very similar argument: ‘It is the fact that I am responsible for what my body does, even when it is not a product of my agency, that gives rise to the permission to harm innocent attackers and innocent threats. And I bear that responsibility because my body *is me*’. Tadros goes on to argue that to fully explain why it is permissible to kill the falling person one needs to appeal to an additional lesser evil justification. I set this aside, since adding this justification is compatible with everything else I have to say.

My second reply to the skeptic is that much of the temptation to connect rights with agency can be explained away. McMahan and others have suggested that there must be a tight connection between infringing a right and exercising agency because we need to explain why creatures that cannot exercise agency cannot infringe rights. For instance, when a tiger attacks me it does not infringe any rights of mine. Is not a person who is not exercising agency, they ask, relevantly similar to a tiger, and thus incapable of infringing rights?

This argument is too quick. We need to distinguish between the conditions a creature must meet to be *the sort of thing* that one can hold rights against, and the conditions under which someone's rights are actually infringed (McMahan 2005, p. 388). It is plausible that (at least one of the) things that make humans, and not tigers, the sort of thing against which one can hold a right is that humans are capable of agency. This is a distinct claim from the one that we are concerned with: that in some cases a right can be infringed even though the person the right is held against is not exercising agency. Still some people will insist that the former claim *implies* the latter. Only humans are subject to moral constraints, they will say, because only humans are capable of understanding and responding to moral norms. And this means that all moral norms must be 'action guiding': they must be the sort of thing that can taken into account by an agent and used by her to regulate her behavior.

How does this affect what we say about rights? We can distinguish between narrow and broad readings of the claim that morality must be 'action guiding'. On a narrow reading, it must be possible at every moment to know and to do what other people's rights demand of us. So interpreted, the claim is clearly too strong. On my way to return a computer on loan from you I run into a completely unexpected traffic jam. I am stuck: there is nothing I can do to get the machine there by the agreed-upon time. This does not make your right to have it back by that time disappear. We can see this by noticing that the right has various implications that go beyond making it the case that, all else being equal, I ought to get your car into the driveway on time. For instance, it makes it the case that I ought to call and warn you that I cannot get the computer there and that I ought to seek a release from you, perhaps offering some payment. And if these things are not possible, or I just do not do them, then the right will require me to compensate you for any losses incurred. Any plausible version of the 'action guiding' claim must be much weaker. Rather than requiring that we can always do what it takes to satisfy rights, it might require that all rights can have some potential influence on our behavior.

This weaker requirement still rules out your having a right against me that, say, it does not rain this weekend—there is nothing I can do about *that*—but it does not rule out the rights that I have been defending. Take the car case discussed earlier: while incapacitated, I can do nothing to remove the car from your drive and thus can do nothing to prevent the infringement of your right against its being there. But that right can still influence my broader behavior, by demanding, for instance, that I pay compensation based on the length of time the car has been there.

Thus, it is plausible to think that a right can be infringed without any exercise of agency. And this means that the rights theory can plausibly explain why defense is permissible even in a case like **Falling Person**, where the threatener is entirely deprived of agency. But, as I said earlier, even if some readers remain convinced

that rights infringement requires the exercise of agency, for my purposes it is enough to notice that the rights theory can explain **Refill** and **Cell Phone**, because, as we will see, the other major theories are not able to explain these cases.²⁰ The rights theory would not be quite able to explain all of our common-sense judgments, but it would still be ahead of the alternatives (at least until someone develops a wholly new theory that can explain all of those cases *plus* **Falling Person**).

The Relevance of Responsibility

We have so far considered cases that suggest a limited role for responsibility in the common-sense morality of self-defense. But there are other cases where the responsibility of a threatener does seem to matter somewhat. I will note two familiar types of example.

First, how much force you may use against a threatener does seem to vary somewhat with her responsibility for the threat she poses. Compare the following two sorts of threat.²¹

- (1) **Drugged:** A threatener is attempting to destroy your prized vase. The reason for her attack, however, is that she is under the influence of a drug which makes it hard (though not impossible) to control oneself. She is under the influence of the drug because a villain secretly put it in her drink earlier in the evening.
- (2) **Clear Mind:** A threatener is attempting to destroy your prized vase (making just the same motions as in **Drugged**). However, she is under no chemical influence and has an entirely clear mind. She simply wants to hurt you.

Even if we agree, as most people do, that you may defend your vase in both of these cases, the maximum amount of force you may use seems to differ. McMahan suggests that you could break a bottle over the head of the threatener in **Clear Mind**, but that doing so would be excessive in **Drugged**. The most you could do is, say, kick her in the shin, even though this may mean the loss of your vase. You need not agree with these precise judgments: just that the maximum amount of harm you can inflict in **Drugged** is *less*.

This much suggests that responsibility plays some role in determining what you may do in cases of self-defense. So it appears that responsibility must affect liability to harm. In particular, it suggests that a threatener's degree of responsibility affects

²⁰ Some people would find it surprising if one could use force in **Cell Phone** but not **Falling Person**, since according to common-sense morality you can kill in both and the attackers appear to be morally indistinguishable, given their similar innocence, lack of responsibility, and likelihood of causing harm. But these cases should only be treated differently on the rights theory if the exercise of agency is essential to rights infringement. And the person who is adamant that rights infringement requires agency has a strong reason to think that there *is* in fact an important morally relevant difference between the two cases—only one involves the exercise of agency—and thus should not be surprised if the two cases call for very different responses on the part of the victim.

²¹ I borrow these examples from Jeff McMahan (2011, Chapt. 4.1.1), who uses them to make the same point about proportionality.

how much harm it would be *proportionate* to inflict on her.²² The less responsible she is for the threat, the less harm it would be proportionate to inflict on her.

The rights theory, at least in the form that I have so far described and defended, appears unable to explain the feature of common-sense morality just noted. On that approach, people become liable to defensive harm when and because they are going to present a wrongful threat. And how much harm they are liable to is fixed by the severity of the wrong. If, as I have argued, responsibility doesn't affect whether I am presenting a wrongful threat, or how severe a wrong I threaten to inflict, then responsibility cannot affect proportionality, or liability. So the rights theory has no explanation for why the threateners in **Drugged** and **Clear Mind** differ in their liability.

In sum, while the rights theory is well placed to explain the limited role for responsibility in self-defense it seemingly has no resources for explaining the ways in which responsibility does seem to affect what we can inflict in self-defense.

Alternatives

How should we respond to all of this? I want to consider whether there is any way to explain all of the features of common-sense morality discussed so far: the ways in which responsibility does and does not seem to matter in the context of defense. The rights theory does not seem to be capable of capturing both. Is there an alternative that is?

The Distributive Justice Theory

The distributive justice theory, which I will now briefly explain and motivate, offers one very plausible way to explain the relevance of responsibility.²³ In self-defense cases, a threat has been created and someone is inevitably going to be harmed. If the initial victim of the threat does nothing, then she will be harmed, and if she decides to defend herself then the threatener will be harmed. So we can plausibly view the moral situation this way: someone is going to be harmed, so who should it be? Should the threatener suffer the harm or the victim? According to the distributive approach we should consider in particular, for whom it would be fair to suffer the harm. For instance, would it be fair for the threatener to suffer the harm or the victim? In the case we considered earlier, where the threatener was engaging in

²² An alternative explanation of these intuitions is that more responsible attackers are generally more determined and thus more difficult to repel. Thus, rather than appealing to proportionality we might appeal to necessity to explain our intuitions: we might say that greater force may typically be used against more responsible attackers because greater force is typically necessary. This suggestion deserves much more extensive treatment than I can pursue here. Very briefly, I think there are some strong intuitions that cannot be explained by this account. For instance, it seems that one can use greater force to stop the culpable person in **Clear Mind** than to stop a person who has been hypnotized to act in the same way, even though the latter will be just as a determined until the hypnosis wears off.

²³ As discussed in McMahan (2005, pp. 394–395).

senseless or malicious violence, it seems clear that it would be fair for her to suffer the harm rather than the victim.

Having determined what distribution of harm would be fair we can then assign *liability* to harm. If it is fair for someone to suffer a harm, then they are liable to having that harm inflicted on them. It is fair for the malicious threatener to suffer the harm she has created, so she becomes liable to having that harm transferred to her. We can ask how exactly this works: why exactly does the fairness of someone's suffering harm change her rights not to be harmed? Plausibly, fairness can play some role at all points in this explanation. It is fair for the malicious threatener to suffer the harm she has created. So it is fair for others to transfer that harm to her. And if it is fair for others to transfer that harm to her then they do not wrong her when they transfer that harm to her: she is liable to being harmed by them in this way.

Now, we have been considering the harm that a malicious threatener is going to inflict and what it would be fair for her to suffer instead. But what if the threatener is not malicious, and going to inflict harm because of a mistake (an understandable mistake, let us say, though one she could have avoided with greater care)? Surely this affects how much harm it would be fair for her to suffer? There is still an important difference between her and her innocent victim: she is more responsible for creating a situation in which someone must inevitably be harmed. But she is less responsible for creating this sort of situation than her malicious counterpart. Thus, the amount of the harm she is required to bear out of fairness is less than what the malicious threatener is required to bear. More generally, we can say that, on the distributive justice approach, responsibility affects proportionality because the amount of harm a threatener is required to bear out of fairness rests on how responsible she is for the threat she poses.

This means that the distributive justice approach can quite naturally accommodate the ways in which responsibility seems relevant to what we are permitted to do in self-defense. For instance, since the threatener in **Clear Mind** is more responsible than the threatener in **Drugged**, it is fair for her to suffer more of the inevitable harm. And so she is liable to that greater degree of harm.

The trouble with the distributive justice theory is that it cannot also accommodate the cases where a lack of responsibility or very limited responsibility does not preclude liability. This is most obvious when we consider threateners who bear no responsibility, such as those in **Cell Phone** and **Falling Person**. According to the distributive justice theory, fairness does not justify shifting harm onto these people and so one cannot use defense force against them. The non-responsible threat is in morally relevant respects just like an innocent bystander, to whom we clearly cannot transfer harm.

But some people may be willing to accept that defense is not permitted against non-responsible attackers. So let us consider again minimally responsible threats, of the kind seen in, say, **Refill**. In these cases, the threateners are performing actions—such as shooting a water gun—that carry *some* risk of harm to others. However, the risk is very slight and so their responsibility for the threat that they pose is at most very slight. Can the distributive justice approach explain why it is permissible to use force, even lethal force, against them?

Given that they are performing activities that carry *some* risk of injury to others, it does seem fair for these threateners to bear *more* harm than their victims. But what if the harm is indivisible? What if, say, the defender's only options in **Refill** are to kill or be killed? It seems permissible for her to kill. And it is hard to see why fairness would demand this.²⁴

First, the difference in responsibility between Threatener and victim in **Refill** is quite limited. Although Threatener has performed an action that poses some risk of harm to others, that risk is extremely low and so she is only a little more responsible for the presence of the threat than the victim.

Second, it might be tempting to say that even if Threatener's responsibility is quite limited it is still enough to *tip the balance* in favor of her bearing the harm rather than the victim. Look, it might be said, ideally Threatener and the victim would bear harm in proportion to their responsibility, but since the harm is indivisible the fairest alternative is for Threatener to bear all of the harm.

This balance tipping argument rests on a mistake. It assumes that the default fair/just allocation of harm is for an innocent threatener and an innocent victim to *equally split* any inevitable harm, and that if the harm is not divisible then fairness/justice permits either to be harmed (or perhaps permits a coin toss or other fair procedure for allocating harm). Only against this background would it make sense to say that a little bit of responsibility on the part of the threatener is sufficient to tip the balance in favor of her bearing the harm. But such neutrality cannot be the default for the distributive justice theorist. For, as explained earlier, distributive justice theorists say that you may *not* transfer harm to a non-responsible threat, even though you are both equally innocent. And distributive justice theorists must say this, because otherwise they would have no explanation for why we cannot transfer harm to, or split harm with, bystanders who are also equally innocent.²⁵

In sum, distributive justice theorists agree, and must agree, that there is a strong presumption in favor of not harming the innocent, including innocent attackers. This means that they cannot explain why self-defense against non-responsible threats is permitted. It also means that they cannot explain the permissibility of defense against a minimally responsible threat: her responsibility is not significant enough for fairness to override the presumption in favor of not harming the innocent.

The Modified Rights Theory

Another option would be to adopt a heavily modified rights theory. The version of the rights theory that I have been discussing has it that responsibility is irrelevant to whether someone is going to infringe a right. One could modify the theory by adopting alternative criteria for rights infringement. According to the modified rights theory (MRT), someone who harms with full responsibility infringes a more significant right

²⁴ McMahan (2011, Chaps. 4.1–4.2) essentially acknowledges this problem.

²⁵ For instance, they might say that even if it would be fair for the innocent victim to share half the harm with an innocent threatener or bystander, a prohibition on doing harm bars the victim from shifting harm to either of these others in the name of fairness. But then it would be very unclear why this prohibition does not also trump the slight increase in fairness of shifting lethal harm from an innocent victim to a minimally responsible threatener.

than someone who inflicts the same amount of harm with lesser responsibility (say, because they have some sort of excuse for inflicting the harm). People may be attracted to this view because it seems to them that someone who inflicts harm on you with full responsibility treats you worse than someone who inflicts harm on you with less responsibility. I think that this apparent attraction rests on a confusion between blameworthiness and wrongdoing: while the former person acts in a more blameworthy way towards you, this does not mean that she is committing a worse wrong. But in this section I just want to consider the implications of accepting the MRT, especially what it has to say about the cases that I presented earlier.

The MRT can readily explain the cases where responsibility appears to have some role to play in fixing what one can do in self-defense. Take the contrast between **Drugged** and **Clear Mind**. On the natural construal of the MRT, the stringency of the right one infringes in harming someone is proportionate to one's responsibility for the harm. The MRT thus entails that the right that the threatener in **Drugged** will infringe if she destroys the vase is less stringent than the right that the threatener in **Clear Mind** will infringe if *she* destroys the vase, given that the latter attacker is more responsible. Hence, the threatener in **Drugged** is liable to less (maximal, necessary) defensive harm, and this fits with our common-sense judgments about how much defensive force you may use against each of the threats.

The problem for the MRT is that, like the distributive justice theory, it seems unable to explain our judgments about the cases discussed where the role of responsibility seems more limited. We saw there that force, including lethal force, may be used against a lethal threatener who is only minimally responsible, or even completely lacking in responsibility. This is very puzzling on the MRT.

We are allowed in self-defense contexts to treat threateners in ways that would not normally be permissible, including, especially, harming them. This is because in presenting threats they come to lack certain rights, including rights not to be harmed. Which rights do they come to lack? Proportionality constraints dictate that the stringency of the rights they come to lack must be in proportion to the stringency of the rights that they threaten to infringe. Now, we saw earlier that it seems permissible to respond to minimally responsible and non-responsible (lethal) threateners by intentionally killing them, which is to say killing them with full responsibility. But according to the MRT, the minimally responsible or non-responsible threatener is going to infringe a much less significant right than the right not to be harmed with full responsibility (given that the stringency of the former rights is proportionate to the threatener's responsibility, and those threateners bear at most a small fraction of the responsibility of a fully responsible threatener). So, the MRT does not predict that one comes to lack the (very stringent) right not to be harmed with full responsibility just by presenting a threat that one is minimally responsible for (or, *a fortiori*, not responsible for): that would be disproportionate. So the MRT does not predict that minimally responsible or non-responsible threateners are liable to being intentionally harmed in self-defense. So, according to the MRT, and contrary to our common-sense judgments, one may not intentionally kill a minimally responsible or non-responsible threatener in self-defense.

It should come as no surprise, then, that modified rights theorists, such as David Rodin (2002, p. 78), recommend that we give up on our common-sense judgments

about the use of force against minimally responsible and non-responsible threats: that is the natural conclusion to draw from their theory. Perhaps there is a way for the modified rights theorist to explain how huge differences in responsibility do not translate into significant differences between the wrong done by someone minimally responsible for harming you and someone fully responsible, but it is hard to see how this can be achieved without making ad hoc claims.

The Rights Theory Supplemented

To present my alternative, I will need to first explain some distinctions. First, note that sometimes even if someone else lacks a right that you not X you still should not X.²⁶ For instance, suppose that someone trespasses on my property and thus I can remove him from the premises without infringing any right of his. It still might be the case that I should let him remain on the premises. For instance, perhaps there are many landmines just outside the premises that are very hard to detect. In that case, considerations other than his rights dictate that I should leave him be. In particular, considerations of beneficence require that I let him stay in the garden rather than face extreme danger. Beneficence sometimes requires that we help those who are in need, even if they have no right to our help and are liable to being treated in ways that would deny them the help.

To be clear, the beneficence that I have in mind here is morally *required*. Sometimes when we speak about beneficence, or charity, we have in mind things that it would be good of me to do or praiseworthy, but which I am not morally required to do.²⁷ For instance, it would be good of me to volunteer to help my busy neighbor weed his garden, but it is not morally required. On other occasions, aiding others is morally demanded. For instance, if while driving through some remote countryside I chance on a hiker by the side of the road who is bleeding heavily, I am required to drive her to the nearest hospital. The duties of beneficence that I have in mind are of this latter kind.

Second, notice that what I owe needy people as a matter of beneficence is sensitive to two other considerations. One, it matters how much the needy person is *responsible* for being in their deprived situation. For instance, if someone is in bad shape because they have been engaging in foolish and reckless driving we are much less willing to help them than if they were damaged in a freak accident. Two, how morally *deserving* they are matters. If someone injures themselves in the course of trying to steal something, we feel a much reduced obligation to help them than if they suffer the same injury while walking innocently down the street. The moral quality of their behavior affects whether we have an obligation of beneficence to them (or how strong that obligation is).

Rights theorists can appeal to these considerations. The theory tells us when and why a threatener becomes *liable* to defensive harm: when and because she presents

²⁶ As Thomson (1986, pp. 68–71).

²⁷ Thomson (1971, pp. 62–63) marks this difference by distinguishing between ‘Good Samaritanism’ and ‘Minimally Decent Samaritanism’.

a wrongful threat. It tells us, for instance, that if a threatener is about to break my leg then she lacks a right that I not, say, break her arm to prevent this (assuming that doing so is the only way to protect myself). It is compatible with this view that other considerations can also play a role in determining whether I *should* (or *may*) now break her arm, all things considered. For instance, perhaps harming her would very badly affect many other people. Obligations of beneficence to threats themselves can also play a role. Even if someone is liable to being harmed by me to a certain degree, I may still be required, out of beneficence, not to inflict that full amount of harm on her.

This allows the rights theorist to explain why there may be a difference between the amount of harm you can inflict on a completely non-responsible threat versus a culpable threat. Compare again the difference between a drugged threatener, who is attempting to smash your prized vase, and a threatener who is trying to inflict the same amount of harm intentionally (as described in **Drugged** and **Clear Mind**, above). People think that the maximum amount of force you may use against the former is less than what you may use against the latter: for instance, lethal force may only be appropriate against the intentional threatener. An explanation for this is on hand. Even if the non-responsible threat is liable to a certain amount of harm, beneficence dictates that you should not inflict the full amount of harm on her.

When you are faced with an innocent threatener you should, out of beneficence, be willing to limit the amount of harm you inflict on her to something below what she is liable to. Even if she is liable to, say, have a bottle broken over her head, you should be willing to risk more damage to your vase in order to inflict less harm on her: just a kick in the shin, say. But obligations of beneficence can be lessened substantially, or entirely, by considerations of responsibility and desert. A culpable threatener is both responsible for being in a situation where it is necessary to harm her and also acting in a morally blameworthy way. These considerations suggest that your beneficence does not require you to limit the amount of harm you inflict on her and so you can inflict the full amount of harm that she is liable to.

Now, common-sense morality suggests that it would not only be praiseworthy to inflict less harm on the less responsible (or non-responsible) but that we are *required* to inflict less harm on them. Can the view that I have proposed really explain this? I suggested earlier that there is a distinction between beneficence that is merely praiseworthy and beneficence that is morally required. So to explain the common-sense view, my account needs to assert that inflicting less harm on less responsible threats is matter of morally required beneficence. Can this assertion be supported?

Here are three reasons to think that it can. First, duties of beneficence are generally heightened where we are dealing with people who have significant needs, and protecting someone from substantial physical harm serves a significant need. Second, duties of beneficence are also generally heightened where one person is uniquely able to help a needy person. For instance, in the hiker example that I used to introduce the category of morally required beneficence you are the only person in a position to help the hiker and this is plausibly part of why you must take her to hospital. Third, duties of beneficence are plausibly heightened where we are considering *doing* harm to someone rather than merely *allowing* harm to befall them. For instance, evicting someone from your property into a dangerous situation

seems worse than failing to help someone in similar danger. In the examples we are considering, what is at stake is how much harm you are going to do to the threatener rather than merely allow her to suffer, and so the duty of beneficence to limit that harm is heightened. In sum, it is plausible to think that a duty of morally required beneficence is present when you are defending yourself against an innocent or less responsible threatener rather than a duty of merely discretionary beneficence. And that is why you are morally obliged to inflict less harm on such a threatener than on a more responsible threatener.

Thus, the rights theory is compatible with the thought that you should inflict less harm on a threatener to the extent that she is less responsible or culpable for the threat she presents. Does it, with the proposed additions, still allow you to inflict lethal harm on minimally responsible and non-responsible threats? Does beneficence require you to give up your life for them? It does not. While the details of obligations of beneficence are debated, it is generally agreed that we are not required to sacrifice our lives or fundamental interests for the sake of others. For instance, beneficence does not require of me that I dive into the ocean, putting myself at very significant risk of death, in order to save someone who is drowning (though it would be highly praiseworthy of me to do so). Similarly, beneficence does not require me to refrain from killing someone who is liable to being killed by me, when doing so is the only way to save myself.

But why is there no difference between the amount of harm I can inflict on a lethal non-responsible threatener and a lethal responsible threatener: why am I permitted to kill them both, if I have stronger obligations of beneficence to the former? The answer is that we do not see any difference because you are allowed to inflict death on the non-responsible threat and there is no further harm available, as the cases are described, that you could additionally inflict on the responsible threatener. If we changed the cases so that the maximum amount of (necessary) harm you could inflict on each was somehow much higher (not just death but lingering death, plus torture, plus...) then we might well reach a point where what you could do to the non-responsible *versus* the responsible would differ.

In sum, we can retain the familiar version of the rights theory, along with its attractive ability to explain why we can harm even non-responsible, or minimally responsible threateners, in self-defense. All we need to do is notice that there are other moral considerations, to do with beneficence, which, in conjunction with the rights theory, allow us to explain our judgments about the cases where responsibility seems to matter. It is not ad hoc to make these additions since we need them anyway to explain other parts of the morality of helping people in need.

Some Objections and Implications

Before concluding, I would like to consider some implications of the theory I have defended and discuss whether those implications are objectionable.

Compensation

I have argued that duties of beneficence sometimes require us to limit the amount of harm that we inflict on a threatener, even if they are liable to the greater degree of harm. In **Drugged**, for instance, you could protect your vase by kicking the threatener in the shin rather than by hitting her over the head: if hitting her is your only option, then you must do nothing. But suppose that you do hit her over the head all the same. It might be said, plausibly, that you should *compensate* her for the additional harm of the clobbering over the shin kick. And yet if my account is correct, then it does not seem compensation is owed, since we do not have to compensate people for our failures of beneficence.

It is true that my account cannot strictly speaking predict that you should compensate the threatener in this situation, because compensation is, by definition, a way of correcting a rights infringement and, on my account, the threatener's rights have not been infringed. However, it can predict that you ought to help the threatener out and give her resources that will reduce the suffering you caused her. This is because even if you do not owe her compensation you still have a duty of beneficence to help innocent people in need: people like the threatener. Moreover, beneficence will typically be morally *required* in this case: her need is significant; you have a special connection to the harm she is suffering, since you inflicted it; and you will likely be in a special position to help her, since she is right in front of you, you know the nature of her injuries, and so on. We cannot technically call this morally required help '*compensation*', but that does not seem to me to be a great loss.

Threatener Self-Defense and Third-Parties

Suppose, again, that in **Drugged** the victim responds to the threatener by smashing her over the head. It might plausibly be said that in this case the threatener herself is allowed to counter-attack (or maybe we should say 'counter-counter-attack!') to protect herself from suffering this harm, and that a third-party could also justifiably use force against the victim to protect the threatener. Yet, on my account, the victim is not infringing anyone's rights by smashing the threatener over the head, since the threatener is liable to the harm that the victim inflicts on her. And so the victim does not become liable to any defensive force that might be used by the threatener or by a third-party. So, my account appears to have a potentially counter-intuitive implication, prohibiting some counter-attacks and third-party defense that seem permissible.

Now, we can sometimes wrong someone in order to produce a sufficiently large benefit for someone else.²⁸ For instance, suppose that I have a pill which is of little value to me (it would only cure a mild headache), but which someone else needs to cure an illness that will otherwise kill them. If I all the same hold on to the pill, it seems permissible for the sick person or for a third-party to steal the pill in order to

²⁸ See Thomson's discussion in *The Realm of Rights*, p. 122, though note that her terminology is different from mine.

save her life. We can justify using force against the victim to *some* degree in this way. For instance, if a third-party could prevent the head-smashing by briefly manhandling the victim, causing only mild bruising, then this might be a justifiable infringement of the victim's rights (including, perhaps, his rights over the vase).

However, this maneuver can only justify fairly mild uses of force against the victim, since to justify wronging someone you need to produce a significantly larger benefit for others. Are threatener and third parties permitted to use greater force than my account would predict? I do not have very clear intuitions about this, though others may. In the end, this may be a place where the theory I have defended requires some alterations to our common-sense picture of when defensive force may be used, as perhaps any account must. My hope is that those revisions are an acceptable price to pay for adopting a view that allows us to explain our central intuitions about self-defense and responsibility.

Conclusion

In sum, if we want to capture our central judgments about responsibility and self-defense we should not replace or modify Thomson's rights theory but supplement it. We can then account for both the role and limits of responsibility in this context. And the resulting account has interesting implications for how we should think about less settled areas of the morality of defense, including self-defense on the part of less responsible threateners and third-party defense.

Acknowledgments For helpful comments, I am very grateful to Tyler Doggett, David Mapel, Jeff McMahan, and two anonymous referees.

References

- Doggett, Tyler. 2011. Recent work on the ethics of self-defense. *Philosophy Compass* 6: 220–233.
- Hosein, Adam. 2014. Are justified attackers a threat to the rights theory of self-defense? In *How we fight*, ed. Helen Frowe and Gerald Lang, 87–103. Oxford: Oxford University Press.
- Kamm, Frances. 1992. *Creation and abortion*. Oxford: Oxford University Press.
- McMahan, Jeff. 2005. The moral basis of liability to defensive killing. *Philosophical Issues* 15: 386–405.
- McMahan, Jeff. 2009. Self-defense against morally innocent threats. In *Criminal law conversations*, eds. Robinson Paul H, Ferzan Kimberly, Garvey Stephen. Oxford: Oxford University Press.
- McMahan, Jeff. 2011. *Killing in war*. New York, NY: Oxford University Press.
- Otsuka, Michael. 1994. Killing the innocent in self-defense. *Philosophy and Public Affairs* 23: 74–94.
- Quong, Jonathan. 2009. Killing in self-defense. *Ethics* 119: 507–537.
- Rodin, David. 2002. *War and self defense*. Oxford: Oxford University Press.
- Tadros, Victor. 2011. *The ends of harm*. Oxford: Oxford University Press.
- Thomas, Scanlon. 2008. *Moral dimensions: Permissibility, meaning, blame*. Cambridge, MA: Belknap Press.
- Thomson, Judith Jarvis. 1971. A defense of abortion. *Philosophy and Public Affairs* 1: 47–66.
- Thomson, Judith Jarvis. 1986. *Rights, restitution, and risk*. Cambridge, MA: Harvard University Press.
- Thomson, Judith Jarvis. 1991. Self-Defense. *Philosophy and Public Affairs* 20: 283–310.
- Thomson, Judith Jarvis. 1992. *The realm of rights*. Cambridge, MA: Harvard University Press.