

Self-Defence and Innocent Threats: The Moral Inequivalence of Innocent Threats and Bystanders

Andreas Sorger

Abstract

Many formulations of a self-defence principle run into significant problems when dealing with cases involving innocent threats since no rights violations have occurred. In response, some philosophers have proposed that killing innocent threats in self-defence is impermissible- what I call the impermissibility thesis. This is based on the notion that innocent threats and bystanders are morally equivalent. This essay argues that there is, in fact, a morally relevant distinction between innocent threats and bystanders such that it is not necessarily impermissible to harm the former. This distinction is based on the difference between eliminative and exploitative harms- where the former is deemed more justifiable than the latter.

*Self-Defence and Innocent Threats: The Moral Inequivalence of
Innocent Threats and Bystanders*

Further, this essay defends this distinction against a recent objection by Michael Otsuka, by arguing for the difference between the right one has to the space one occupies, and the entitlements that come with occupying a space.

The concept of self-defence, although almost universally shared, runs into significant problems when considering cases that involve innocent threats. In response, some philosophers have proposed that killing innocent threats in self-defence is impermissible- what I call the impermissibility thesis. This is based on the notion that innocent threats and bystanders are *on a par*; that is, they are morally equivalent. Contra this, I argue that there this, in fact, a morally relevant difference between innocent threats and bystanders such that it is not necessarily impermissible to harm the former. This difference is explained by the distinction between eliminative and exploitative harms, and shows the impermissibility thesis to rest on a false assumption. I also make the further claim that any self-defence principle must have a decision procedure in repeat cases in order to defend it from what I call the *repeated threat objection*. I, therefore, conclude that proponents of the impermissibility thesis cannot defend their conclusion on the basis that innocent threats and bystanders are on a par.

I proceed in four parts: first, I present a principle of self-defence and use Nozick's Well Case to highlight the contrasting intuitions between the perceived permissibility of killing an innocent threat and

the impermissibility of killing an innocent bystander. Specifically, I highlight the problems surrounding the application of Thomson's self-defence principle in the Well Case as no rights violations occur. Second, I present Otsuka's argument that it is impermissible to kill an innocent threat, before presenting my critique that there is a morally relevant distinction between innocent threats and bystanders. Third, I consider and respond to objections based on the impermissibility of eliminative killings in certain cases, as well as the strength of one's right over the space one occupies, if this space is publicly owned. Fourth, I consider the implications of this distinction in repeat instances of innocent threats, and show how this can lead to moral issues- namely the killing of too many people.

At this point, it is important to note that all the cases I consider are *one-on-one* cases. More precisely, unless otherwise specified, the principles of self-defence elucidated here will solely focus on instances where one innocent threat is falling and threatening the life of one other person, or repeated instances of one threat, innocent or otherwise, threatening the life of one other person. This is to remove utilitarian considerations of the *greater number*, as, intuitively, even those most committed to views that rights are entirely inviolable would be willing to concede that it would be permissible to violate the rights of one individual in order to save the lives of a significant number of others- for instance a million. Hence, in order to remove all considerations of number, I follow the trend in the literature by only considering instances where one life is at risk.

1 Innocent Threats and the *Well Case*

An initially plausible principle of self-defence can be formally phrased as follows:

B can justifiably act in self-defence towards an aggressor A if and only if A has violated, or threatens to violate, B's rights.

However, this principle does not capture the distinction between aggressors and threats, where the former involves an individual A actively seeking to harm a different individual B while the latter merely involves A to unintentionally pose a threat to B. For instance, A would be an aggressor if he were to chase after B with an axe while possessing the conscious intent to harm if not kill B, whereas A would be a threat to B if he is pointing a very real-looking gun at B without B knowing it is a paintball gun. In both instances, it would seem appropriate to claim that B can respond in self-defence as A continues to violate B's rights when A is a threat, giving the following principle:

*B can justifiably act in self-defence towards an aggressor OR threat A if and only if A has violated, or threatens to violate, B's rights.*¹

This is similar to the self-defence principle proposed by Thomson, and while this may initially appear uncontroversial, it is only so in instances where A is a culpable threat or aggressor. That is, it has its

¹ Judith Jarvis Thomson, "Self-Defense." *Philosophy & Public Affairs* 20, no. 4 (1991).

strongest intuitive force where A knowingly and willingly threatens B's life and, therefore, violates B's rights. However, could such a principle apply in instances where A knowingly but unintentionally threatens B's life such that B's rights were not necessarily violated? Consider the following: suppose Jack finds himself, through no fault of his own, trapped at the bottom of a well with a ray gun as he was fetching a pail of water. Suppose further that some evil villain Jill picks up a third party Jane and throws her into the well such that if she were to land on Jack he would die while she would survive. Would it be permissible here for Jack to fire his ray gun and vaporise Jane before she lands on him? ²

This proves to be a problem for this principle of self-defence as it is not clear who is violating Jack's rights. Clearly, Jill violates both Jane's and Jack's rights by throwing Jane into the well, since she is threatening both of their lives, yet the key debate centres around whether or not Jane violates Jack's rights if she were to land on him. For Thomson, the mere fact that Jane threatens Jack's life, even if she does so involuntarily, is sufficient to claim that Jane violates Jack's right to life, and therefore allows Jack to justifiably respond in self-defence by vaporising her. Yet, it is not altogether clear whether this is, in fact, true; Jane was forcibly thrown into the well by Jill, and therefore bears no responsibility for her current situation, which in turn arguably absolves her of any rights violations.

² Robert Nozick, *Anarchy, State, and Utopia*, Malden, MA: Blackwell, (2012): 34-35.

Here, Jane is an *innocent threat*; she is a threat because she directly threatens Jack's life, yet she is also innocent because she is not responsible for being thrown in a well and hence it would be counterintuitive to claim that she is violating Jack's rights. Consider the following analogous case: suppose Jack is sitting at the base of a cliff. Suppose further that a large rock, through the natural process of erosion and without human interference, falls from the cliff such that it lands on Jack, killing him. Here it would be ludicrous to claim that the rock has violated Jack's rights, as this would lead to the absurd conclusion that anyone who died at the hands of an inanimate object had their rights violated by the object. ³Yet, there are striking parallels between the rock and Jane; both threaten Jack's life and both have no form of agency, thereby suggesting that Jane does not, in fact, violate Jack's right to life.

2 Otsuka and the Moral Equivalence Thesis

From this, Otsuka argues that innocent threats are on a par with innocent bystanders, or innocent third parties that have no claim to the issue at hand, and it would, therefore, be impermissible for Jack to shoot his ray gun.⁴ This is based on the strong intuition we possess that killing innocent bystanders in self-defence is completely impermissible. Note that Otsuka is making three claims: he claims

³ Michael Otsuka, "Killing the Innocent in Self-Defense," *Philosophy & Public Affairs* 23, no. 1 (1994): 80.

⁴ *Ibid.*, 82-84.

that it is impermissible to kill an innocent bystander in self-defence, and from this argues that, as innocent threats are on a par with bystanders (the moral equivalence thesis), it is impermissible to kill an innocent threat in self-defence (the impermissibility thesis)⁵. That is, because there is the strong intuition that one should not treat bystanders as a *mere means*, one cannot permissibly kill an innocent bystander in order to achieve some pre-defined goal. So, it would be as impermissible to use a bystander as a human shield against a projectile as it would be for Jack to kill Jane since Otsuka views both scenarios as being indistinguishable. This is because both Jane and the bystander share an important morally relevant quality- namely, they have no responsible agency. Therefore, under Otsuka's account Jane is a bystander in a *morally relevant sense* as, while she does threaten Jack's life, this is without a desire or volition to do so- thereby making her morally equivalent to a bystander⁶.

I find this view to be mistaken on an important count- namely that there is a *morally relevant* distinction between bystanders and innocent threats. Consider the following amendment to the well case (I call this the Gun Case): suppose Jack was in possession of a normal gun that he could use to shoot and injure Jane, thereby slowing her momentum such that she will land on Jack and just injure, rather than kill him. Contrast this with the 'Mattress Case', in which Jill throws Jane off the roof of a building, where she will land on Jack unless he shoots and injures a bystander holding a mattress, such

⁵ Ibid.,76.

⁶ Ibid., 84.

*Self-Defence and Innocent Threats: The Moral Inequivalence of
Innocent Threats and Bystanders*

that the mattress will give Jane a soft landing⁷. Under Otsuka's account, it would be impermissible for Jack to injure both Jane in the Gun Case and the bystander in the Mattress Case as this would violate their right to not be harmed when neither has violated any of Jack's rights; in fact, it would be akin to me injuring pedestrians with a baseball bat. However, I contend that it is intuitively permissible for Jack to injure Jane in the Gun Case, purely on the basis that Jane threatens Jack's life, but not to injure the bystander in the Mattress Case as the bystander has no relevance to the causal chain. More precisely, Jack treats the bystander as a mere means in the Mattress Case, yet intuitively *does not* treat Jane in the same manner in the Gun Case.

This distinction can be elucidated using Quinn's distinction between eliminative and exploitative harms. More precisely, the difference between an innocent threat and a bystander can be explained by the permissibility of eliminative harms vis--vis exploitative harms. The former involves simply removing the obstacle the victim presents, whereas the latter involves one 'benefitting from his presence'⁸, or using him opportunistically for the purpose of achieving some goal. Here, an exploitative harm seems *worse* than merely an eliminative one, as while the victim is harmed in both senses, there is the added intuition that they are also *used* in the former case- thereby mak-

⁷ Note that this intuition shift is not due to the fact that Jack is not injured in the Mattress case; if the example is amended to reflect this, the intuition that injuring the bystander is wrong remains.

⁸ Warren S Quinn, "Actions, Intentions, and Consequences," *Philosophy & Public Affairs*, (1989): 344.

ing it harder to justify. Applying this distinction to the two cases outlined earlier, it is clear that Jack harms Jane eliminatively in the falling case, yet harms the bystander exploitatively in the Mattress Case. This undermines Otsuka's view that innocent threats are morally equivalent to bystanders, as this implies either that eliminative and exploitative harms are equally hard to justify, which I have shown not to be the case, or that there is a difference in justifiability, but despite this, they remain equally impermissible. Yet neither implication matches with our intuitions in the above cases, and therefore Quinn's distinction not only explains our intuition shift but also shows that it is not necessarily impermissible to fire on innocent threats in self-defence.

3 Objections and Responses

One could object on the grounds that it is not necessarily the case that eliminative harms are easier to justify than exploitative harms, even in cases when one's life is at risk. Consider the following case: suppose that Liam is trapped underground with a train coming straight for him. Suppose further that there is an alcove he can run into that is deep enough for one person to stand in, yet that this alcove is already occupied by another person, Brad⁹. If Liam were to pull Brad out of the alcove such that the train would hit him instead and Liam would be safe, then Liam is harming Brad eliminatively.

⁹ Jonathan Quong, "Killing in Self-Defense," *Ethics* 119, no.3 (Apr., 2009): 526

*Self-Defence and Innocent Threats: The Moral Inequivalence of
Innocent Threats and Bystanders*

This is because Liam is not using Brad to achieve some goal; in fact, if Brad were not there, Liam would be safe as he could hide in the alcove. As he is not benefitting from Brad's presence, any harm that Liam inflicts on Brad is necessarily eliminative.

Yet, here it seems intuitively impermissible for Liam to throw Brad in front of the oncoming train. Thus a proponent of the impermissibility thesis could raise the following objection: as it is impermissible for Liam to kill Brad in self-defence, and that killing Brad is morally equivalent to killing Jane since in each case the harm inflicted is eliminative, then it follows that it must be impermissible to kill Jane. This is because neither Liam nor Jack benefit from either Brad's or Jane's presence; rather, they would both prefer if the other person was not there, as then there would either be an avenue to escape a life-threatening situation (Alcove Case), or there would be no threat to one's life (Gun Case). Note that even if the example is changed to reflect the Gun Case (suppose instead of killing Brad, that Liam could somehow throw him into the path of the oncoming train such that he is badly injured), the intuition that this is impermissible does not change. Therefore, not all cases of eliminative harm are permissible, and if it is deemed impermissible to harm Brad in the alcove case, then it is impermissible to harm Jane.

In response, I argue that there is an important difference between the Alcove and Gun cases that has not been considered- namely who has the right to occupy that space. This response is similar to the one put forward by Jonathan Quong and holds that if individuals

occupy a certain space, then they have a prior claim to being in that space¹⁰. In the Alcove Case, Brad has a legitimate claim to the space he is occupying inside the alcove, in virtue of standing in it, and therefore if Liam were to throw him in front of the train, he would be completely ignoring Brad's claim to space he is currently occupying favour of ensuring his own survival. This would be a complete violation of Brad's strongly held rights; in fact, to paraphrase Quong, Liam would be using Brad's 'entitlements against him' (i.e. the space he occupies) for his own survival, and therefore under his account Liam would be using Brad as a 'mere means', which makes this action impermissible¹¹. Consequently, by appealing to the idea of prior claims, the shift in intuition in the alcove case can be explained by the fact that it is an exploitative harm.

In a recent article, Otsuka has offered a response, claiming that while individuals have a right to the space they occupy, this claim is not strong enough to justify the impermissibility of killing in the Alcove Case. More precisely, Otsuka argues that an individual's claim to the rights over the space they occupy does not apply if the space they are occupying happens to be a public place, since one lacks property rights over the space that is being claimed, and thus makes any claim unconvincing¹². Consider the following example: suppose Betty is sitting on a park bench such that she occupies the only spot

¹⁰ Ibid., 528.

¹¹ Ibid., 528-30.

¹² Michael Otsuka, "The Moral-Responsibility Account of Liability to Defensive Killing," In *The Ethics of Self-Defense*, ed. Christian Coons and Michael Weber, 51-68. Oxford University Press, 2016: 60.

in the entire park with a view of the lake¹³. Suppose further that she occupies this space from ‘sunup to sundown’, and when asked to move by Albert, a third party, she refuses and claims that ‘by getting there first’, she has a ‘rightful claim to this spot’¹⁴. This reply is entirely unconvincing; the park bench is not hers, and it would only be fair for her to share the view with other people, as is intended in virtue of being public property. That is, as the bench is public property, enjoyment of the bench (and therefore the view) must be shared with all, as no one individual has any prior claim to it.

In fact, Otsuka goes as far as to argue that, if the bench were positioned on a Lazy Susan, it would be permissible for Albert to rotate the bench, without Betty’s permission, such that she would be moved from her spot and he can subsequently enjoy the view¹⁵. This has implications for the alcove case; as the alcove is in a public subway, Brad does not own the alcove he is standing in and therefore has no prior claim to it in virtue of ‘standing in it first’. At best, Brad has an *equal claim* to the alcove as Liam, since neither owns the alcove, and thus it would be justifiable for Liam to unilaterally impose a coin toss on Brad, and subsequently claim the alcove if he wins¹⁶. This is especially true in cases where the alcove is on a Lazy Susan where Liam can rotate Brad out of the alcove such that he is in the path of the oncoming train since this would not violate Brad’s

¹³ Ibid.

¹⁴ Ibid.

¹⁵ Ibid.

¹⁶ Ibid., 61-62.

self-ownership rights¹⁷. Consequently, this undermines Quong's argument for the impermissibility of killing Brad, as, based on claims over the right he has to the space he occupies, Brad has at best an equal claim to the alcove as Liam.

This is a powerful objection that has not yet been responded to, yet I would like to offer a potential route for a response. I would claim that there is a morally relevant distinction between the right over the space one occupies and the right to the entitlements that come with occupying a particular space. Consider the following: suppose that I am second in a queue for a street-food vendor and that I am behind Alice who is taking her time. Suppose further that she is on a Lazy Susan such that, if I rotate it, she will end up behind me as second in line. Here, I think it would be intuitively impermissible for me to rotate the Lazy Susan, *no matter how long she takes*, as she has the right to occupy the first space in the queue in virtue of her being there. Compare this to Otsuka's bench case, where it is intuitively permissible to rotate Betty away from the view. In both cases, an individual is occupying a public space, and on Otsuka's account neither have a claim to the space they occupy as it cannot be grounded in a fundamental right such as property. So what explains this intuition shift?

I argue that this can be explained by the distinction between a right over the space one occupies, and a right to the entitlements of the space one is occupying. This is because, in public places, no one has

¹⁷ *Ibid.*, 61.

*Self-Defence and Innocent Threats: The Moral Inequivalence of
Innocent Threats and Bystanders*

a right to any of the entitlements offered by any space, irrespective of whether or not it is occupied. That is, in the bench case, nobody is entitled to have the view of the lake since this is a good that should be shared with everyone; it is simply the case that anyone occupying that particular spot on the bench can *benefit* from having this good rather than having a right to it. But, anyone has the right to occupy any *free* space in the public area such that, if asked to move, it would be within one's rights to refuse. So, Albert could ask Betty to move from the bench, but she has a legitimate right to say 'no'- even if this may be unconvincing. However, it would be within Albert's rights to benefit from the entitlements of the lake view if he could do so without violating Betty's right over her spot on the bench since she has no claim over the view. This, therefore, explains the intuition shift as such an outcome can happen in Otsuka's bench case- namely if Albert rotated the bench by 180 degrees. Here, Betty's right to the space she occupies is not violated, as she is still seated on the same end of the bench, and Albert can now benefit from the lake view- something Betty did not have a right to. Further, in the queue case, such an outcome cannot happen as the Lazy Susan would *move* Alice from being first in line to second in line- thereby violating her right to the space she occupies. By extension, as rotating Brad in the alcove case would see him be placed in front of the train, this would be impermissible as Liam would be using his right to his space 'against him'- thereby 'shifting the blame of death onto him'¹⁸ which would be impermissible.

¹⁸ Quong, "Killing in Self-Defense," 528-29.

4 The ‘Repeated Threat’ Objection

The ‘repeated threat’ objection holds that, if it is justifiable in one case to harm an innocent threat, then does it follow that it is justifiable in all cases? Consider the following: suppose an evil villain constructs a machine such that it selects a bystander at random and throws them into the well, where Jack is once again trapped at the bottom through no fault of his own and in possession of a ray gun. Further, suppose that the machine continues to throw bystanders for the duration of an hour (or until Jack dies). As I have argued previously, it is permissible for Jack to harm the innocent threat, but it seems problematic to say that Jack can continue to harm each subsequent innocent threat, as this only constitutes an eliminative harm, and is therefore permissible when one’s life is at risk.

I contend that the answer to this revolves around the need for a decision procedure in repeated threat cases. Susanne Burri has argued that one should flip a coin when considering whether or not to harm an innocent threat, as this is the only way a fair distribution of harms can come about¹⁹. However, she would hold that this still applies in one-person cases, which would undermine my argument for the permissibility of killing a threat without a coin toss in such cases. In response, I would argue that the intuition behind my argument stems from the moral permission we have to prioritise our own lives over others; as we only have one life, we have permission

¹⁹ Susan Burri, “The Toss-Up Between a Profiting, Innocent Threat and His Victim,” *Journal of Political Philosophy* 23, no. 2 (Jun., 2015).

*Self-Defence and Innocent Threats: The Moral Inequivalence of
Innocent Threats and Bystanders*

to protect it to some extent at the expense of saving others²⁰. It is this priority that allows one to permissibly refrain from saving a drowning person if one is an awful swimmer, as if one has a high chance of dying in saving the child, it would be permissible for one to not save the person drowning. Note that this moral permission is also compatible with rights as side constraints since morality does not require us to sacrifice ourselves in order to save others.

In repeated threat cases, however, I would argue that this notion of moral permission for the prioritisation of our own lives does not apply. This is because in such a case, one would have to kill numerous individuals in order to save one's life and therefore one is arguably not prioritising one's life, but rather *valuing* it above all others, which I would argue is impermissible. For instance, while it would be permissible to kill an innocent threat and save our life, it would not be permissible to blow up a train full of people to save our life. Therefore, as the harm must be distributed amongst multiple individuals, a decision procedure, such as a coin toss, must be employed in order to ensure that the harms are justly distributed.

To conclude, I argued that there is a morally relevant difference between innocent threats and bystanders such that it is not necessarily impermissible to harm the former. This is based on a distinction between eliminative and exploitative harms, where the former is deemed more justifiable than the latter as it does not involve the use of an individual as a mere means, and that this distinction can be

²⁰ Quong, "Killing in Self-Defense," 516-17.

Andreas Sorger

sustained when faced with objections. I also made the further claim that, in repeated cases, a decision procedure is required in order to ensure the just distribution of harms across multiple individuals, which is not required in single threat cases due to the permissibility of prioritising one's own life.

References

- [1] Burri, Susanne. “The Toss-Up Between a Profiting, Innocent Threat and His Victim: Profiting Innocent Threat and His Victim.” *Journal of Political Philosophy* 23, no. 2 (Jun., 2015).
- [2] Nozick, Robert. *Anarchy, State, and Utopia*. Malden, MA: Blackwell, (2012).
- [3] Otsuka, Michael. “Kamm on the Morality of Killing.” *Ethics* 108, no. 1 (Oct., 1997): 197-207.
- [4] Otsuka, Michael. “Killing the Innocent in Self-Defense” *Philosophy & Public Affairs* 23, no. 1 (1994): 74-94.
- [5] Otsuka, Michael. “The Moral-Responsibility Account of Liability to Defensive Killing.” In *The Ethics of Self-Defense*, ed. Coons, Christian., and Michael Weber, 51-68. Oxford University Press, (2016).
- [6] Quinn, Warren S. “Actions, Intentions, and Consequences: The Doctrine of Double Effect.” *Philosophy & Public Affairs*, (1989): 334-351.
- [7] Quong, Jonathan. “Killing in Self-Defense.” *Ethics* 119, no. 3 (Apr., 2009): 507-37.
- [8] Thomson, Judith Jarvis. “Self-Defense.” *Philosophy & Public Affairs* 20, no. 4 (1991): 283-310.

Andreas Sorger

Andreas Sorger is a Politics and Philosophy graduate. His main fields of interest are Normative and Applied Ethics, alongside the Philosophy of Public Policy. He can be contacted at [a.j.sorger@lse.ac.uk]