

The Resurrection of Radical Pacifism: A Defense

Stephen Blake Hereth

A dissertation

submitted in partial fulfillment of the  
requirements for the degree of

Doctor of Philosophy

University of Washington

2019

Reading Committee:

Michael I. Blake, Chair

Stephen M. Gardiner

Colin Marshall

Program Authorized to Offer Degree:

Philosophy

© Copyright 2019  
Stephen Blake Hereth

University of Washington

**Abstract**

The Resurrection of Radical Pacifism: A Defense

Stephen Blake Hereth

Chair of the Supervisory Committee:

Michael I. Blake

Department of Philosophy

Within the ethics of self-defense, the predominant view is that there are liability justifications for harming. A minority position, which I call radical pacifism, denies that there are liability justifications for harming. This dissertation offers three separate arguments against the predominant view and for the radical pacifist view. The first paper, “Animal Rights Pacifism,” demonstrates how attentiveness to our moral duties to animals generates counterintuitive moral conclusions that can be most plausibly avoided only by appeal to pacifism. The second paper, “Multiple Threats and the Specter of Pacifism,” shows that anti-pacifism is itself radical, since it entails that there can be *unlimited* liability justifications for harming. The final paper, “Against Self-Defense,” makes use of Nozickian moral side-constraints and defends the existence of the Cavalier Constraint according to which killing ought never to be done with moral ease. The anti-pacifist view allows for the violation of this constraint, however, and is therefore false.

## Introduction: The Papers, the Methods, and Future Projects

### 1. The Three Papers

Aside from a few corners of philosophy, pacifism—the view that there are no liability justifications for harming—is dead, or at least rotting.<sup>1</sup> This dissertation involves three related papers, each of which seeks to resurrect pacifism as a live ethical and political position. I shall now briefly describe *how* the three papers accomplish this task.

The first paper, “Animal Rights Pacifism,” illustrates how speciesism can be used to ground considerations leading to a defense of pacifism. As the first paper demonstrates, pacifism—as a virtually eliminativist view of defensive permissions—possesses explanatory power about how we should treat those who harm nonhuman animals. Those who defend animal rights and non-rights-based duties toward animals are largely anti-pacifist, yet they often find themselves without a plausible explanation for the impermissibility of animal rights militancy.<sup>2</sup> The argument defended in the first paper therefore shows that pacifism has *explanatory power* that anti-pacifism arguably lacks over cases like these. Indeed, while some may argue that speciesism hides the permissibility of animal rights militancy from us, it may also be that speciesism disaffects our moral intuitions by considering the permissibility of self-defense *in isolation from* cases involving defense of nonhuman animals. That is, it may be that the widespread acceptance of an anti-pacifist view of defensive permissions leaves so-called ‘marginal’ cases unconsidered, but that a lack of

---

<sup>1</sup> For example, Cheyney Ryan’s defense of radical pacifism continues largely unabated. See Ryan, “Pacifism, Just War, and Self-Defense,” *Philosophia*, Vol. 41, No. 4 (2013): 1-29. Cf. Saba Bazargan, “Varieties of Contingent Pacifism in War,” in *How We Fight: Ethics in War*, edited by Helen Frowe and Gerald Lang (NY: Oxford University Press, 2014), 1-17.

<sup>2</sup> Tom Regan, for example, claimed that pacifism was obviously false, but not (as Jan Narveson maintained) logically incoherent. See Regan, “A Defense of Pacifism,” *Canadian Journal of Philosophy*, Vol. 2, No. 1 (1972): 73-86.

consideration of these cases conceals important moral insights about the *impermissibility* of self-defense. This epistemology connects tightly with a major insight within philosophy of disability: namely, that if our ethics fails to consider the value and rights of disabled individuals, then we run a serious risk of endorsing false ethical theories (or false variations of them), false moral principles, or misguided intuitions about particular cases.<sup>3</sup>

Another important conclusion reached in the first paper is that pacifism must reject liability justifications for defensive harming. In the debate over moral liability, it is possible for pacifists to maintain that although some individuals have forfeited their right against harm (including, perhaps, lethal harm), this alone is insufficient to ground the permissibility of killing them. On such a pacifist view, the impermissibility of defensive killing is therefore not grounded in a rejection of moral liability, but of moral liability *as a justification*. This is structurally similar to Jeffrey Reiman's retributivist argument that although perpetrators of serious crimes deserve to be executed, it's nevertheless impermissible to execute them.<sup>4</sup> But it is also possible for pacifists to claim that even individuals who are culpably responsible for unjust harms to others don't forfeit their right against harm (including lethal harm). On this view, such individuals *maintain* their right against harm. This has the effect of drawing a battle line for future defenses of pacifism. Such a commitment is also demonstrated in the third paper, "Against Self-Defense," where a side constraint of permissible killing is grounded in an individual's right against being cavalierly killed. Since this right entails a duty not to kill, killing therefore violates that right, at least presumptively.

The second paper, "Multiple Threats and the Specter of Pacifism," demonstrates how a novel problem for a particular view of moral liability—the so-called Problem of Multiple Threats

---

<sup>3</sup> For an argument to this effect, see Joseph A. Stramondo, "Why Bioethics Needs a Disability Moral Psychology," *Hastings Center Report*, Vol. 46, No. 3 (2016): 22-30.

<sup>4</sup> See Jeffrey Reiman, "Why the Death Penalty Should Be Abolished in America," in *The Death Penalty: For and Against*, Jeffrey Reiman and Louis P. Pojman (Oxford: Rowman & Littlefield, 1998), 67-132.

(hereafter, PMT)—also implicates a broader anti-pacifist view of moral liability. The general worry behind the PMT is that an anti-pacifist view of moral liability allows for excessive killing. This has obvious political ramifications within the ethics of war: When states fight wars, they kill numerous people. If there is some plausible upper limit on the number of culpable aggressors one can permissibly kill, this seriously restricts warfare. This shows that the PMT is problematic not only for the dominant view of moral liability *within* the anti-pacifist camp, but that it's also problematic for *all* anti-pacifist views of moral liability. (Indeed, it implicates modern warfare itself.) Since pacifism does not suffer from the PMT, this offers reason to reject anti-pacifist views of moral liability. If a pacifist view is the only alternative (and I argue that it is), this indicates that we should embrace the pacifist view of moral liability.

The third paper, "Against Self-Defense," shows that a Nozickian conception of side-constraints on permissible harming can be used to defend pacifism. In the paper, I defend a particular side-constraint: namely, that even in cases where killing is morally permissible, one ought never to kill cavalierly (i.e., casually). This principle, which I call the Cavalier Constraint, is widely endorsed among self-defense ethicists. The general idea is that because human life has considerable value even when human rights are forfeited, one's decision to kill ought to be appropriately considered and undertaken with moral trepidation. However, I argue that since the permissibility of defensive violence can be radically overdetermined, and since the extent of our moral deliberation and moral trepidation ought to be responsive to moral reasons, this would justify us in killing cavalierly in some cases. Thus, the anti-pacifist view that violence *can* be justified itself violates the Cavalier Constraint, since it allows for the kind of overdetermination that would justify cavalier killing. This essay demonstrates that pacifism upholds a high view of the moral value of persons. Their value is in fact so high that they possess a *moral right* to have their lives

and welfares regarded with the utmost moral seriousness. The existence of this right prohibits act-types that could, in principle, violate that moral right.

Each of these papers is a small attempt to justify a resurgent pacifist research project. Although perhaps not decisive, each paper makes headway in showing that pacifism is indeed a live ethical and political position. As I describe in section 3, major findings within these three papers lay the groundwork for the next stage of pro-pacifist research.

## **2. Methods**

Some might question the philosophical methodology used in the three papers. However, I believe that this objection does not succeed, for reasons I shall now address. It's worth noting at the outset that while I stand by the methodology in the dissertation, I also believe that my *usage* of the methodology is compatible with those who reject it. In what follows, I shall attempt to explain (1) what my methodology is, (2) why I believe it is fundamentally sound, and (3) why it's worth using for the three papers even if unsound.

Philosophical methodology is contentious, including within ethics. There are those who, on the one hand, prefer what I shall call an *intuitionist* approach to doing ethics. Intuitions, as I use the term, as intellectual seemings: an intellectual indicative sense that some proposition is true or false. These are distinct from perceptual seemings, which originate from the senses. My sense that  $2+4=4$ , for example, is an intellectual seeming, whereas my sense that it is warm outside is an example of a sensory seeming. Michael Huemer, a defender of ethical intuitionism, explicitly makes this comparison:

Appearances can be intellectual, as opposed to sensory, mnemonic, or introspective. It seems to us that the shortest path between any two points must be a straight line; that time is one-dimensional and totally ordered (for any two moments in time, one is earlier than the other); and that no object can be completely red and completely blue at the same time. I accept those things on intellectual grounds. I am not looking at all the possible pairs of points and all the possible paths connecting each pair and seeing, with my eyes, that the straightest path is the shortest in each case. Instead, I am ‘seeing’ intellectually that it must be true—that is, when I think about it, it becomes obvious.<sup>5</sup>

However, these seemings have similarities.<sup>6</sup> For example, they are both *indicative* of a proposition’s truth or falsity. This is true irrespective of their reliability. If it seems to me that it is warm outside, that is an indication that it is warm outside even if it is a false indication. Similarly, the fact that something seems impermissible to me is an indication that it is indeed impermissible even if it is in fact permissible. Provisionally, then, we can say, following Blake McAllister, that intuitions are “*sui generis* mental states with propositional content and a distinct phenomenal character.”<sup>7</sup>

Intuitionists, however, don’t simply believe that intuitions are a phenomenologically special kind of thing. They also hold that they are an *epistemically* special kind of thing, in that that they offer (presumptive) positive epistemic status for beliefs. Here’s one (admittedly internalist) way to characterize intuitionism: The view that intuitions are *reasons for belief*. To say that intuitions are reasons for belief is to say that the presence of an intuition, such as *it seems true*

---

<sup>5</sup> Michael Huemer, *Ethical Intuitionism* (NY: Palgrave MacMillan, 2008), at 100.

<sup>6</sup> See, for example, Elijah Chudnoff, “Intuition in Mathematics,” in *Rational Intuition: Philosophical Roots, Scientific Investigations*, edited by Barbara Held and Lisa Osbeck (NY: Cambridge University Press, 2014), 174-191.

<sup>7</sup> Blake McAllister, “Seemings as *Sui Generis*,” *Synthese*, Vol. 195, No. 7 (2018): 3079-3096.

*that all bachelors are unmarried males*, offers presumptive justification for a belief. Thus, if I have an intuition that *p* is true, and if I lack any good reason to believe that *p* is false, then I am all-things-considered (albeit weakly) epistemically justified in believing *p*. Intuitions, therefore, are evidence.<sup>8</sup>

As a view, intuitionism is disputed. Both historically and recently, there have been intensive efforts to undermine intuitionism. Some of these efforts are led by skeptics, but others maintain that intuitionism is a rusty tool best replaced with a more reliable philosophical methodology.<sup>9</sup> For example, Herman Cappelen argues that intuitions are a lazy trademark of philosophy and that while philosophers often *think* they are relying on intuitions, they are relying on nothing of the sort.<sup>10</sup> If Cappelen is right, then a common argument for intuitionism—that intuitions are indispensable for most knowledge—is significantly weakened.<sup>11</sup> Others have argued that philosophers are indeed relying on intellectual seemings, but that these seemings are unreliable or very narrowly reliable.<sup>12</sup>

While I can scarcely hope to solve this dispute, it's worth pointing out that these criticisms are far from decisive. Indeed, they have all encountered considerable pushback.<sup>13</sup> Within the last

---

<sup>8</sup> For more on this view, see Michael Huemer, "Phenomenal Conservatism and the Internalist Intuition," *American Philosophical Quarterly*, Vol. 43, No. 2 (2006) 147-158. Cf. William G. Lycan, "Phenomenal Conservatism and the Principle of Credulity," in *Seemings and Justification: New Essays on Dogmatism and Phenomenal Conservatism*, edited by Chris Tucker (NY: Oxford University Press, 2013), 293-305; Richard Feldman and Earl Conee, *Evidentialism* (NY: Oxford University Press, 2004), at 15.

<sup>9</sup> For a broad attack on intuitionism, see Michael Raymond DePaul and William M. Ramsey (eds.), *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry* (Lanham, MD: Rowman and Littlefield, 1998).

<sup>10</sup> Herman Cappelen, *Philosophy Without Intuitions* (NY: Oxford University Press, 2014).

<sup>11</sup> For a reply to Cappelen on behalf of intuitionism, see David Chalmers, "Intuitions in Philosophy: A Minimal Defense," *Philosophical Studies*, Vol. 171, No. 3 (2014): 535-544.

<sup>12</sup> See, for example, Eric Schulz, Edward T. Cokely, and Adam Feltz, "Persistent Bias in Expert Judgments About Free Will and Moral Responsibility: A Test of the Expertise Defense," *Consciousness and Cognition*, Vol. 20, No. 4 (2011): 1722-1731.

<sup>13</sup> See, for example, Ernest Sosa, "A Defense of the Use of Intuitions in Philosophy," in *Stich and His Critics*, edited by Michael Bishop and Dominic Murphy (Malden, MA: Blackwell, 2009), 101-112.

decade, a ‘Rationalist Renaissance’ has undergirded a renewed defense of intuitionism.<sup>14</sup> Criticisms of intuitionism within psychology and experimental philosophy have encountered criticism of their own. Matthew Liao, for example, argues that empirical studies in fact *demonstrate* the reliability of ‘robust’ intuitions, and thus fail to undermine intuitionism generally.<sup>15</sup> Jennifer Nagel argues (persuasively, in my view) that empirically-motivated anti-intuitionist arguments often show merely that our intuitions are fallible, but that this fails to undermine the general reliability of our intuitions unless we are prepared to say that our sensory seemings (such as perception) are also not generally reliable.<sup>16</sup>

There are three reasons I shall proceed as if intuitionism is true: two substantive and one provisional. The first substantive reason that I assume intuitionism is for the simple reason that I believe it’s true. While I don’t offer a thoroughgoing defense of intuitionism, I see little reason to suppose that I am not entitled to assume its truth. Doubtless other ethicists proceed as if certain theses are true despite the fact that I deny them, and anyway each of us must assume *some* things if we are to make philosophical progress. This would be less defensible if our weighted evidence disconfirmed intuitionism. However, I am unconvinced of this claim and am joined by a chorus of voices in my skepticism. The second substantive reason is because intuitionism is a useful tool for parsing our beliefs. Intuitionism clarifies *why* we believe as we do, why we are reluctant to embrace certain views and arguments, why we avoid certain actions (e.g., because they seem impermissible), and so on. Anti-pacifist views arguably rely as much on intuitionism as pacifist views, and perhaps even more so. Thus, a pacifist examination of our moral intuitions can be useful

---

<sup>14</sup> See, for example, Andrew Chapman, Addison Ellis, Robert Hanna, Henry Pickford, and Tyler Hildebrand, *In Defense of Intuitions: A New Rationalist Manifesto* (London: Palgrave MacMillan, 2013).

<sup>15</sup> S. Matthew Liao, “A Defense of Intuitions,” *Philosophical Studies*, Vol. 140, No. 2 (2008): 247-262.

<sup>16</sup> Jennifer Nagel, “Intuitions and Experiments: A Defense of the Case Method in Epistemology,” *Philosophy and Phenomenological Research*, Vol. 85, No. 3 (2012): 495-527.

in clarifying why anti-pacifists have the intuitions they do, what those intuitions actually are, and whether an anti-pacifist view of defensive permissions fundamentally *depends upon* the assumed reliability of certain intuitions. If this is true, then even pacifists who *reject* intuitionism can use an examination of it as a philosophical cudgel against anti-pacifists. Third, and finally, intuitionism is largely *assumed* in most of the literature on the ethics of self-defense, and a key assumption is that the pacifist position must be rejected if commonsense morality is to be maintained. It seems fair to question this assumption on its merits: *Is* anti-pacifism the most intuitive position? The dissertation below contains three major arguments against that view. Moreover, *if* intuitionism is true—a possibility that should be conceded even by anti-intuitionists—it would hardly be fair to deny pacifists an opportunity to marshal a defense of their view using intuitions.

### **3. Future Projects**

This dissertation advances three major arguments for pacifism. The first is an argument grounded in the rights of many nonhuman animals. The second focuses on a new problem in the literature, the Problem of Multiple Threats. The third defends a side constraint to permissible killing that vindicates pacifism. These arguments, however, are part of a broader project to resurrect a pacifist view of defensive permissions.

The dissertation does not give all that is owed, however. Both pacifists and anti-pacifists are, in my view, owed a more comprehensive vision of a pacifist ethic. This is especially true since pacifism, if practiced consistently, would upend many of our current practices and institutions. It is therefore a radical moral position that requires not merely considerable defense, but flesh on the

bones. Despite these limits, the defenses of pacifism herein nevertheless lay the groundwork for a more exhaustive defense and articulation of the position.

The conclusions reached in the three papers, described in section 1, establish fertile soil in which to plant further pro-pacifist projects. To begin, a pacifist view of the state would arguably exclude liability justifications for other forms of defensive or punitive harming such as arrest, imprisonment, and restrictions on migration. It remains less than obvious whether a pacifist could condone these harms for other reasons even if the individuals' right against them remains intact. Nor is it clear whether pacifists can endorse procreation, where serious harms are reasonable foreseeable, or defensive killing in which (like euthanasia) the victim is made better-off by being killed. The temptation, of course, is to simply invoke a consent-based justification for these actions. However, fetuses cannot consent at the time of their creation, and the limits of retroactive consent, hypothetical consent, and tacit endorsement are far from clear.<sup>17</sup> Nor is it clear that consent will excuse euthanasia in a way that leaves pacifism unscathed since some, like Thomas Hurka, argue that the principal justification for the defensive killing of combatants is that they consent to be killed in virtue of voluntarily undertaking their combatant status.<sup>18</sup> While this shows that the anti-pacifist 'plant' project is yet ill-watered, it also shows that a more comprehensive pacifist project has scarcely begun to sprout. The very distinctions that underlie much of anti-pacifist moral theory must be revisited if pacifism is to be believed. To use a somewhat Cartesian metaphor: If the plant above is poisoned, what of the ground below it?

Yet another thing owed is a full account of what pacifism is, both in terms of what it prohibits and requires and what it permits. For example, does pacifism require radical duties of

---

<sup>17</sup> See, for example, Asheel Singh, "The Hypothetical Consent Objection to Anti-Natalism," *Ethical Theory and Moral Practice*, Vol. 21, No. 5 (2018): 1135-1150.

<sup>18</sup> Thomas Hurka, "Liability and Just Cause," *Ethics and International Affairs*, Vol. 20 (2007): 199-218.

aid? Relatedly, what is a pacifist theory of healthcare justice? Does pacifism require a rejection of burdensomeness as a mitigating factor for moral obligation? Finally, what common deontological moral distinctions can pacifists consistently accept? A more thorough pacifist vision, both of the state and of a reconceived normative ethics, clarifies both where pacifist has taken us and where it *should* take us. The present dissertation therefore represents a beginning for a broader project and lays the groundwork for future theory.

## **Animal Rights Pacifism**

**ABSTRACT:** The Animal Rights Thesis (ART) entails that nonhuman animals like pigs and cows have moral rights, including rights not to be unjustly harmed. If ART is true, it appears to imply the permissibility of killing ranchers, farmers, and zookeepers in defense of animals who will otherwise be unjustly killed. This is the Militancy Objection (MO) to ART. I consider four replies to MO and reject three of them. First, MO fails because animals lack rights, or lack rights of sufficient strength to justify other-defensive killing. Second, MO fails because those who unjustly threaten animals aren't liable or, if they are liable, their liability is outweighed by other considerations (e.g., a strong presumption against vigilante killing). I then argue both of these fail. Third, MO succeeds because animal militancy is permissible. Fourth, MO fails because there aren't liability justifications for defensive killing in general (i.e., pacifism is true). I argue that there's thoroughgoing epistemic parity between the Militancy View (MV) and the Pacifist View (PV), and that two considerations favor PV over MV. First, because under conditions of uncertainty, we should believe rights-bearers retain rather than lose their rights, which PV affirms and MV denies. Second, because PV is intrinsically likelier than MV to be true since PV at worst affirms wrongful letting die and MV at worst affirms wrongful killing, the latter of which is intrinsically harder to justify than the former.

### **1. Introduction**

I shall explore the Militancy Objection to the view that certain animals have basic moral rights of comparative strength to the rights we typically assume most human beings have. To explore such an objection, however, it is necessary first to explain what sort of view the objection targets. The Animal Rights View, as I shall call it, consists of a central thesis, which is this:

ANIMAL RIGHTS THESIS:

Many animals have basic moral rights of comparable strength to the basic moral rights possessed by humans.

I'll assume that most farm animals (e.g., dogs, pigs, chickens, cows, and horses) and wild animals (e.g., giraffes, zebras, rhinos, and lions) have these rights while remaining neutral about insects and other animals.

What rights do these animals have? On the minimalist view of animal rights spelled out in the ANIMAL RIGHTS THESIS, animals have basic rights that are comparable to the basic rights we typically assume most human beings have. These include, for example, a right against being unjustly killed, a right against being unjustly tortured, and a right against being killed merely in order to be consumed recreationally (i.e., unnecessarily). Further, because these basic rights are of *comparable strength* to those possessed by most human beings, they entail similar permissions and prohibitions. For instance, it would be a violation of human being's basic rights to kill them for sport, and thus it's also a violation of an animal's basic rights to kill them for sport. Similarly, it would be a rights violation to torture a (non-consenting) human being merely to satisfy research curiosity or to benefit others, or to kill and consume a human being for nourishment or gustatory pleasure when doing so is unnecessary—that is, when there are other reasonably available sources of nourishment. By implication, these actions would also be rights violations if taken against animals. I therefore assume the following thesis:

HARM THESIS:

All animals have a presumptive right not to be harmed.

Perhaps one of the most widely accepted rights is the right to self-defense. This extends to self-defense against animals, as well. It is widely assumed that it is permissible for human beings to kill animals who pose a threat to them. Even strong defenders of the ANIMAL RIGHTS THESIS, such as Tom Regan and Gary Francione, concede that the defensive killing of animals is often

permissible.<sup>19</sup> But as Cheryl Abbate rightly observes, such a position is insufficiently attentive to the broad range of cases:

An ethicist who takes seriously the rights of animals should remain troubled by the seemingly unreflective responses offered by the dominant theories of animal rights regarding the issue of defensive killing of nonhuman animals. This is because when we consider the standard scenarios in which nonhuman animals are said to threaten the lives of human beings, what immediately comes to mind are situations where the lives of humans are threatened *because* of their unjust actions and treatment of nonhuman animals.<sup>20</sup>

Abbate's point in this passage concerns *liability* to defensive harm—that is, the conditions under which one party doesn't wrong another party by defensively harming or killing them. For the animal rights theorist, a plausible liability story will explain not only why animals are sometimes liable to be killed by humans, but also why humans are sometimes liable to be killed by animals. Since the view that humans can't be (or never are) liable to defensive harm is *prima facie* incompatible with the ANIMAL RIGHTS THESIS whereas its negation is *prima facie* compatible, I shall assume that humans are liable to defensive harms by animals when humans unjustly harm animals or attempt to do so. Plausibly, the HARM THESIS implies the following claim:

ASSISTANCE THESIS:

All animals have a right to defensive assistance.

---

<sup>19</sup> See Tom Regan. (1983) *The Case for Animal Rights*. Berkeley: University of California Press, p.296; Gary L. Francione. (2000) *Introduction to Animal Rights: Your Child or the Dog?* Philadelphia: Temple University Press.

<sup>20</sup> Cheryl Abbate. (2015) "The Search for Liability in the Defensive Killing of Nonhuman Animals." *Social Theory and Practice* 41 (1): 106-130, pp.106-107.

Defensive assistance is what it seems like: assistance in the form of defense. Here, some third party is doing the work that the defending party has a right to do—or, in the case of animals, *would* have a right to do if they were moral agents—namely, self-defense. Thus, other-defense is simply assisted self-defense. There is some reason to believe the ASSISTANCE THESIS. Our duties to assist others depend upon whether those individuals are being unjustly harmed in that case.<sup>21</sup> Murderers, for example, lack a right to self-defense against law enforcement attempting to imprison or injure them, but this is not true of murderers' intended victims. If I am right that the ASSISTANCE THESIS follows from the HARM THESIS, then we should expect that defensive assistance of the murderer would not be morally required whereas it would be required for his intended victims. Such a view gets the intuitive judgments precisely right here.<sup>22</sup> Moreover, on the account of rights I am assuming, if you have a duty to treat someone in a particular way, such that the treatment is owed to *that* individual, then that individual has a right to that treatment. Notice also that the right to defensive assistance is, like the right to self-defense, a presumptive right. I might lack a right to *your* defensive assistance in a particular case if defensively assisting me in that case would invariably result in your death or grievous injury. Similarly, if animals have a right against being harmed, their right is a presumptive right and therefore subject to certain restrictions.<sup>23</sup>

---

<sup>21</sup> It's commonly assumed that a *moral agent's* right against (unjust) harm entails a right to self-defense which entails a right to defensive assistance. While it doesn't strictly follow, it would be surprising if the right against (unjust) harm didn't *itself* entail a right to defensive assistance, and thus the presence or absence of moral agency on the part of the victim of unjust harm doesn't seem to matter. For an extended defense of the entailment relation for moral agents, see Suzanne Uniacke. (1994) *Permissible Killing: The Self-Defense Justification of Homicide*. New York: Cambridge University Press, pp.177-179. For a contrary view, see David Rodin. (2002) *War and Self-Defense*. New York: Oxford University Press, pp.17-34 and 38. Rodin's view is that duties to defend others derive not from their right to self-defense, but from more general duties to protect things of value. If Rodin is right, the defender of the ANIMAL RIGHTS THESIS can deny (P3) in the argument below. However, another argument can be easily reconstructed in which the ASSISTANCE THESIS does not derive from the HARM THESIS but nevertheless derives from the ANIMAL RIGHTS THESIS.

<sup>22</sup> We need not accept this strong thesis to generate the Militancy Objection, however. It is sufficient that animals' right against harm entails a presumptive permission for third parties to defensively assist animals.

<sup>23</sup> I make no effort here to say just what these restrictions might be.

But perhaps it's too much to assume that the ASSISTANCE THESIS is implied by the HARM THESIS, or even to assume that the ASSISTANCE THESIS is true. I believe both that it's true and that it follows from the HARM THESIS, but for present purposes it is sufficient to show that an animal's right against harm implies a presumptive *permission* to assist them defensively.<sup>24</sup> Thus, we need only assume that the following thesis is true if the HARM THESIS is true:

PERMISSION THESIS:

Third parties are permitted to defend animals under threat of unjust harm.

---

<sup>24</sup> Two objections are sometimes given to the ASSISTANCE THESIS. The first is that the thesis falsely implies that third parties are *obligated* to assist even if doing so would come at an unreasonable cost to them such as death or severe injury, or perhaps an unreasonable *risk* of unreasonable cost to them. The second is that the thesis falsely implies that we have a right to use the bodies of others, since having a right to third-party assistance entails a right to use their bodies for our defense.

In reply to the first objection, it's crucial to note that the right to defensive assistance is a *presumptive* right, and thus certain considerations (like unreasonable cost) can mitigate its status as a right for third parties to whom the relevant considerations apply.

My reply to the second objection is twofold. First, the objection appears to imply that we have *no* duty to assist others in any case whatsoever, whether they be drowning children or animals under threat of unjust harm by human aggressors. But that's surely false. For a creative and extended defense of this position, see Cécile Fabre. (2006) *Whose Body Is It Anyway? Justice and the Integrity of the Person*. New York: Oxford University Press, chapter 2. Fabre contends that bodies are resources subject to distributive justice and that the unjustly imperiled have a presumptive right to our assistance and the entailed use of our bodies. Similarly, Adam Hosein argues that it's permissible to use the bodies of third parties if doing so is necessary to rescue someone from extreme harm. See Adam Hosein. (2014) "Are Justified Aggressors a Threat to the Rights Theory of Self-Defense?" In *How We Fight*, ed. Helen Frowe and Gerald Lang. New York: Oxford University Press, 87-103, pp.92-93.

Second, it is fairly uncontroversial that we are permitted to exercise at least some control over the bodies of individuals who are liable to defensive harm from us. If individuals who can reasonably assist us and fail to do so are liable to defensive harm as a result, therefore, we are permitted to exercise some control over their bodies. I conclude, tentatively, that such individuals *are* liable to defensive harm from us. Here's why: If you fail to help someone under threat of unjust harm *H* when you could reasonably do so (i.e., when it's feasible and at little cost to you), then you share responsibility for the fact that *H* isn't no longer a threat to that person, which entails that you are responsible for the fact that *H* is *still a threat* to that person, which entails that you bear some responsibility for the fact that the threat imposed by *H* is *maintained*. Since the threatened person has a reasonable claim against you that you not maintain threats, you are liable to defensive harm if you fail to help them when you reasonably could. Moreover, since you can't be liable for acting permissibly and since every act is either permissible or impermissible, it follows that you acted impermissibly. Thus, you have an obligation to assist them resulting from a claim they have on you—which is identical to a right to assistance.

This thesis is compatible with the ASSISTANCE THESIS but is weaker than it, since it claims only that a right to self-defense entails that third party defense is *permissible*, not that it is obligatory (as the ASSISTANCE THESIS implies). Moreover, as in the previous two theses, the PERMISSION THESIS comes with a caveat: The third-party permission to defend the animal is presumptive. Thus, the PERMISSION THESIS doesn't imply that *any* third party is permitted to defensively assist an animal with defensive rights. For example, suppose that Maggie the mutt is wrongly being tortured by Helen the human and that another human, Christine, can defensively assist Maggie. If Christine does so, she will fail to save her own child, and she is *obligated* to save her own child. Because Christine has a prior obligation to save her own child over other individuals with an equally strong moral status, she is *forbidden* from defending Maggie.<sup>25</sup> Whether one finds this persuasive, the point is simply that the PERMISSION THESIS allows for the possibility that some third parties are not permitted to engage in defensive assistance. Hereafter, then, I shall assume that the PERMISSION THESIS is true and that it follows from the HARM THESIS.

## 2. The Militancy Objection

What I have shown thus far is that the Animal Rights View being assumed here has strong implications for moral life. If animals have basic rights of *equal* strength as human beings, then they too have a right not to be unjustly killed. Thus, there is a strong moral presumption that they should not be captured and killed. Under those circumstances, third parties are presumptively permitted—indeed, obligated—to *rescue* them. Now consider

---

<sup>25</sup> For a similar example of an agent-relative duty, see Derek Parfit, (2011) *On What Matters: Volume 1*. New York: Oxford University Press, pp.140-141.

### *Burning Barn*

Sadist wants to burn alive ten innocent pigs. He finds a barn, enters it, and sets it ablaze.

But a wooden beam collapses onto Sadist, preventing him from escaping the barn. Passerby has a choice: save the ten pigs, or save Sadist.<sup>26</sup>

Following Jeff McMahan, it is an *infringement* of rights when utilitarian considerations make it permissible to do what others have a right that we not do, and a *violation* when no considerations (utilitarians or otherwise) make it permissible.<sup>27</sup> Consider a case in which you have a right that I not enter your property without your consent, but I will die unless I step onto your lawn to avoid being hit by a car. There, your right continues to exist despite the permissibility of me working against it, making it unlike cases of violation (in which I work against your right, but not permissibly) and forfeiture or waiving (in which you lack the right).

Sadist's actions violate the pigs' right not to be harmed. If the ANIMAL RIGHTS THESIS is true, it implies that Passerby is permitted to do what they'd ordinarily be permitted to do in a case with human victims: to save the victims (whether humans or pigs) over Sadist. This judgment is intuitive: Sadist acts wrongly, the pigs don't, and this makes a difference about whose rescue morality favors or requires.<sup>28</sup> Indeed, it's intuitive *regardless* of the truth of the ANIMAL RIGHTS THESIS. This is because even if you think the duty to rescue Sadist is presumptively stronger than the duty to rescue the pigs, that's changed by Sadist's unjust actions against the pigs. In other words, Sadist *forfeits* his priority—if not his right to be rescued altogether—because he unjustly

---

<sup>26</sup> My thanks to Helen Frowe for this example.

<sup>27</sup> Jeff McMahan. (2011) *Killing in War*. New York: Oxford University Press, p.10. The third type of case, of course, is one that is neither an infringement nor a violation of someone's right(s). For example, individuals who are liable to be harmed lack a right not to be harmed.

<sup>28</sup> There's something presumptively wrong with saving Sadist over his victims.

threatens the pigs. Because Sadist forfeits this priority, and because Passerby wouldn't wrong Sadist if she chose to save the pigs, there's a *liability justification* for saving the pigs over Sadist in *Burning Barn*.<sup>29</sup>

So far, we've considered a liability justification for not saving humans over animals. It appears there are sometimes such justifications. That itself doesn't seem problematic. Consider a different example:

### *Night Shift Zookeeper*

Nour is a 17-year-old working a summer job at the Toledo Zoo. To reduce the costs of caring for the zoo's perfectly healthy pigs, they will be killed and fed to the other half. Nour is tasked with euthanizing the pigs tonight—the same night animal activists sneak into the zoo to save the pigs from slaughter. If Nour sees the activists, she will call the police and the activists' purpose will be foiled (i.e., the animals will be killed). The activists know this. The only way the activists can save the animals is by surprising Nour and slitting her throat.

It's hard to see how mere cost saving justifies killing the pigs. Certainly it wouldn't if the ANIMAL RIGHTS THESIS were true. But nor would it under less radical moral beliefs.<sup>30</sup> If, for example, I could save a few bucks this month by killing my dog, it would still be wrong of me to do so. Nour's actions, therefore, are unjustified, and so her threat against the pigs is unjust. Just as it did with Sadist in *Burning Barn*, Nour's unjust threatening incurs moral liability upon her. The moral

---

<sup>29</sup> This doesn't mean there's a liability justification for leaving Sadist to die in a case unlike *Burning Barn* where it's possible to save the pigs *and* Sadist.

<sup>30</sup> For an extended defense of this view, see Crystal Allen Gunasekera, "The Ethics of Killing 'Surplus' Zoo Animals," *Journal of Animal Ethics*, Vol. 8, No. 1 (2018): 93-102.

liability in this case isn't about saving, however, for Nour isn't in danger. It's about *killing*, and defensive killing at that: Nour's death is (practically) necessary to save the pigs she unjustly threatens; killing her is instrumental in saving them. When someone is liable to be defensively harmed, there's a defeasible justification for harming them. In *Night Shift Zookeeper*, that justification appears undefeated, and thus it's actually permissible to kill Nour. And if the ASSISTANCE THESIS is true—if the presumptive obligation to save the pigs is undefeated, as it appears to be—killing Nour is *obligatory*. Yet both the claim that it's obligatory to kill Nour and the claim that it's merely permissible to kill her are anathema to our moral intuitions. More formally,

*The Militancy Objection to the ANIMAL RIGHTS THESIS*

- (P1) If many animals have basic moral rights of comparable strength to basic human rights, then many animals have a right against harm. [Assumption]
- (P2) If many animals have a right against harm, then many animals are such that third parties are permitted to defensively assist them. [Assumption]
- (P3) If many animals are such that third parties are permitted to defensively assist them, then the pigs in the Toledo Zoo are such that third parties are permitted to defend them against Nour. [Assumption]
- (C1) Therefore, if many animals have basic moral rights of comparable strength to basic human rights, then the pigs in the Toledo Zoo are such that third parties are permitted to defend them against Nour. [From P1, P2, and P3]

- (P4) If the pigs in the Toledo Zoo are such that third parties are permitted to defend them against Nour, then harming Nour in defense of the pigs in the Toledo Zoo is presumptively permissible. [Assumption]
- (P5) But harming Nour in defense of the pigs in the Toledo Zoo is not presumptively permissible. [Assumption]
- (C2) Therefore, it's not the case that many animals have basic moral rights of comparable strength to basic human rights. [From C3, P4, and P5]

The argument is a *reductio ad absurdum* against the ANIMAL RIGHTS THESIS when wedded to certain plausible auxiliary assumptions. The conclusion, the objection goes, is plainly incompatible with our commonsense moral beliefs about the impermissibility of harming or killing not just about animal researchers, but zookeepers, hunters, combatants in warfare who kill or intentionally endanger innocent animals,<sup>31</sup> and perhaps even poachers.<sup>32</sup> If the ANIMAL RIGHTS THESIS is true, it is actually permissible to harm (and perhaps even kill) many of these individuals—a deeply counterintuitive implication.<sup>33</sup> As John Hadley observes,

[I]f people who are harming animals are liable to third-party defense, then tens of thousands, possibly millions, of well-intentioned, law-abiding, good-natured, talented and

---

<sup>31</sup> Think, for example, of naval personnel who utilize dolphins to identify mines, or ground forces which utilize canines and other animals to detect explosives. Similar arrangements are made by police personnel (e.g., K-9 units).

<sup>32</sup> I say “perhaps” since poachers are less likely to be viewed favorably under commonsense moral views.

<sup>33</sup> Abbate considers a case in which a human man hikes on a wilderness trail where grizzlies are known to roam. He comes across a grizzly, the grizzly growls, and the man fatally shoots the grizzly. Abbate describes this case as one in which the hiker engaged in and is responsible for risk-taking activities which resulted in a foreseeable violent conflict. Because the risk-imposition is non-reciprocal—that is, because “the hiker could have stayed home” and “the bear cannot be expected to just ‘stay in his den’”—the hiker is liable to defensive harm to an extent that the grizzly is not. Thus, the hiker, and not the grizzly, should bear the brunt of the ensuing harm. See Cheryl Abbate, “The Search for Liability in the Defensive Killing of Nonhuman Animals,” pp.122-124.

otherwise reasonable people will be legitimate targets for violence. Presumably, this would mean that many farmers who raise animals for food, scientists who use animals in biomedical research, people from varying occupations who employ animals in entertainment and recreational pursuits, and doctors who abort sentient fetuses, would be liable to third-party defensive violence on behalf of the animals.<sup>34</sup>

That the ANIMAL RIGHTS THESIS appears to imply that these individuals can be permissibly harmed on these grounds is, therefore, an objection against the ANIMAL RIGHTS THESIS.<sup>35</sup> Such an objection is the Militancy Objection.<sup>36</sup>

### 3. Four Responses

How might defenders of the ANIMAL RIGHTS THESIS reply to the Militancy Objection? I shall explore four possibilities and conclude that two are either implausible or not open to defenders of

---

<sup>34</sup> John Hadley. (2009) "Animal Rights and Self-Defense Theory." *Journal of Value Inquiry* 43: 165-177, p.168. For more on a Militancy-Objection-style argument for the permissibility of abortion, see Christopher Tollefsen. (1997) "Donagan, Abortion, and Civil Rebellion." *Public Affairs Quarterly* 11 (3): 303-312. See also Christopher Kaczor. (2015) *The Ethics of Abortion: Women's Rights, Human Life, and the Question of Justice* (Second Edition). New York: Routledge, p.220.

<sup>35</sup> Cf. Rainer Ebert and Tibor R. Machan. (2012) "Innocent Threats and the Moral Problem of Carnivorous Animals." *Journal of Applied Philosophy* 29 (2): 146-159. Ebert and Machan focus on a distinct objection, which they call the Predation Objection. According to this objection, the ANIMAL RIGHTS THESIS implies that agents have a duty to defend animals from predation by other animals, which they claim is absurd. In the course of defending this objection, they note that since Regan *denies* that there is a duty to defend animals from such predation on the grounds that animals are 'moral patients' and therefore cannot violate rights, it follows that it is impermissible to harm or kill animals in defense of humans. They then claim such a view also implies that if, for example, a wolf attacks a human and the human responds with violent self-defense, it is *permissible* to harm or kill the human in defense of the wolf's rights. This is notable for two reasons. First, it shows that the ANIMAL RIGHTS THESIS presumptively justifies animal rights militancy. Second, it portrays this implication as problematic.

<sup>36</sup> See also Jeff McMahan's objection to Judith Jarvis Thomson in Jeff McMahan. (2002) *The Ethics of Killing: Problems at the Margins of Life*. New York: Oxford University Press, pp.398-421. Thomson assumes for argument's sake that fetuses have strong rights and argues that abortion is nevertheless often permissible. McMahan replies that such a strong view of rights implies that a third party can permissibly defend fetuses under immediate threat of abortion.

the ANIMAL RIGHTS THESIS (or both), and that one of the other views must be true if the ANIMAL RIGHTS THESIS is true. I then argue that the final view is true.

### **3.1. Deny or Weaken the Animal Rights Thesis**

The first option is simply to concede the Militancy Objection and deny the ANIMAL RIGHTS THESIS. Because such a move is clearly not open to defenders of the ANIMAL RIGHTS THESIS, said defenders should look to the alternatives.

Moreover, this view depends upon certain speciesist assumptions that do not withstand scrutiny. Suppose that the individuals in the Toledo Zoo were not pigs but adult human beings. In that case, Nour's actions would be morally akin to those of a prison executioner, and it would hardly seem implausible to suppose that she could be permissibly killed to defend her actual or would-be victims. Yet the only difference here is mere species, and a difference in mere species cannot itself make a difference as to whether individuals have rights.

Moreover, even if it did make a difference to whether animals have rights, it would not make a difference to whether animals have a moral status sufficiently robust to justify defensive assistance. If I come across a stranger sexually molesting and torturing an orca, it seems that I am permitted to stop the stranger from doing this even if that requires harming the stranger. This implies that even a rejection of the ANIMAL RIGHTS THESIS is insufficient to evade the force of the problem.

Moreover, if the objection is simply that animals lack rights because, if they had rights, third parties would be permitted (in the course of defensively assisting animals) to inflict immense harm or even death to a vast number of human individuals, then this is prima facie reason to reject

any ascription of rights whatsoever. Consider that history contains various periods in which an enormous percentage of the human population had an abusive hand in the horrors of slavery, often threatening to harm slaves, or actually harming them, or killing them. Ascribing rights to slaves, then, would permit at least as many human beings to be harmed or killed, but that is hardly a good reason to deny that slaves have rights.<sup>37</sup> Thus, we should abandon this response to the argument.

Maybe there's a way to maintain animal rights *and* condemn animal rights militancy. This can't be done by denying animal rights, but that's just as well, since denying animal rights appears to be a weak strategy. What might do the trick is *weakening* the animal rights position. Before we explore ways of doing that, let's first recall the animal rights position we're working with:

ANIMAL RIGHTS THESIS:

Many animals have basic moral rights of comparable strength to the basic moral rights possessed by humans.

The basic moral right at issue with animal rights militancy is a moral right against harm, including harmful death. In the *Night Shift Zookeeper*, militants kill Nour because she unjustly threatens the lives of pigs. Keeping fixed the (plausible) premise that these otherwise healthy, happy pigs have a moral interest in continuing to live and avoiding the physical pain of being killed by Nour, we can infer that Nour *wrongs* the pigs if she causes them pain and ends their lives. This fact alone doesn't entail that Nour is liable to be killed for doing so. After all, killing Nour has to meet several other conditions to be liable to be killed by the militants. For starters, killing Nour must be

---

<sup>37</sup> Speciesists will doubtlessly defend an asymmetry here, contending that human slaves are importantly different from animals, such that ascriptions of rights are not undermined by the permitting of widespread violence in the case of slaves but are undermined in the case of animals. But that's just to offer a distinct objection to the ANIMAL RIGHTS THESIS.

*narrowly proportionate*.<sup>38</sup> This means, roughly, that the harm imposed on Nour shouldn't exceed the harm she threatens. Let's represent this as follows:

NARROW PROPORTIONALITY THESIS:

Killing Threat in defense of Victim's life is narrowly proportionate if and only if Threat's death wouldn't be morally worse than Victim's death.

Applied to Nour in *Night Shift Zookeeper*, this means killing Nour is narrowly proportionate if and only if Nour's death wouldn't be morally worse than the pigs' deaths. And here's the thing: Many philosophers deny that the death of a pig is as bad, morally speaking, as the death of someone like Nour. In fact, many philosophers friendly to animal rights deny it. For example, Jeff McMahan claims that the badness of someone's death depends partly on how strong their moral interests are over time, and that the time-relative interests of most animals are significantly weaker than most humans. Referring to his Time-Relative Interest Account of the wrongness of killing, McMahan writes,

It implies, for example, that the killing of an animal is normally substantially less seriously wrong than the killing of a person [and most human beings are persons].... [The] Time-Relative Interest Account accepts that it is directly relevant to the evaluation of killing that the amount of good an animal loses by dying is typically much less than the good a person loses. Because [prudential unity self-relations over time] are typically weaker in the case

---

<sup>38</sup> I assume internalism about defensive liability merely for the sake of argument. According to internalism, Threat is liable to be harmed only if harming Threat would be necessary to avert some harm, narrowly proportionate, etc. In other words, these requirements are 'internal' to liability; someone isn't liable without meeting the requirements. For more on this distinction, see Frowe, *Defensive Killing*, at 88-89 and 91-4.

of an animal, the Time-Relative Interest Account implies that there is a further reason for discounting the degree to which the killing of an animal is objectionable.<sup>39</sup>

McMahan's talk of "prudential unity relations" over time refers to how an individual's future is tied to their current interests. For example, Nour's current interests are significantly tied to her future: college, partnerships, and career. As a result, if the militants kill her, she loses a great deal; many of her interests are frustrated. How do pigs compare? Let's assume for the moment that pigs' time-relative interests are comparable to dogs. McMahan says this of dogs' time-relative interests:

[It] seems that even adult human life tends to contain its share of exuberant joys that rival in intensity those experienced by dogs. They are simply not so conspicuous as they are within the lives of dogs, where they dramatically punctuate days otherwise given over to torpor and sleep. Human well-being, by contrast, is more continuous, dense, and varied, so that the ecstatic moments, which may be more diffusely spread over longer periods, are less salient. And what fills the intervals between these moments is normally altogether better than the dull vacancy of a dog at rest.... In general, therefore, the quantity of life that an animal loses in dying is less than that which a person loses.<sup>40</sup>

Let's assume McMahan is right that dogs lose less by dying than humans do, at least on average, and thus that the same goes for pigs. Mark Rowlands, another defender of animal rights, concurs with McMahan. Again comparing dogs' interests to humans' interests, Rowlands writes,

---

<sup>39</sup> McMahan, *The Ethics of Killing*, at 194-5. Cf. Jeff McMahan, "The Comparative Badness for Animals of Suffering and Death," in *The Ethics of Killing Animals*, ed. Tatjana Višak and Robert Garner (New York: Oxford University Press, 2016), 65-85.

<sup>40</sup> McMahan, *The Ethics of Killing*, at 196.

[It] is difficult to see that dogs orient much of their present behavior, and discipline many of their present desires, on the basis of their conception of how they would like their future to be. And, if this is right, then the typical dog has far less invested in its future than the typical human, who does orient her behavior and discipline her desires on the basis of such a conception.... Therefore, it would...be irrational to save the life of a dog [over the life of a human].<sup>41</sup>

Rowland's claim is that it's presumptively impermissible to *save* a dog's life over a human's. That generalizes to *killing* a human to save a dog. Given our prior assumption that the loss of a dog's life is equally bad as the loss of a pig's life, it follows that killing a human to save a pig is presumptively impermissible. Molly Gardner, in the course of defending her Attenuated Rights View, makes claims similar to McMahan and Rowlands. After commenting that humans typically live longer than dogs, she adds:

More importantly, even a typical year in the life of a normal human contains more value for the human than a typical year in the life of a normal dog contains for the dog. It is this difference in the severity of the harm [of death], rather than a difference in the species of the one who would be harmed, that grounds the difference in moral judgments.<sup>42</sup>

---

<sup>41</sup> Rowlands, *Animals Like Us*, at 88 and 89.

<sup>42</sup> Molly Gardner, "The Interspecies Killing Problem," in *The Moral Rights of Animals*, ed. Mylan Engel Jr. and Gary Lynn Comstock (Lanham, MD: Lexington, 2016), 119-140, at 129.

Each of these philosophers claims that dogs lose less in death than humans do. By assumption, then, so do pigs. Moreover, since this claim, if true, would apply to many kinds of animals, let's assume that the following thesis is true:

COMPARATIVE HARM THESIS:

An animal's death is itself morally less bad than a human's death.

With these assumptions in hand, we have the makings of an argument against the ANIMAL RIGHTS THESIS. If true, this will undermine (P3) in *The Militancy Objection to the ANIMAL RIGHTS THESIS*. Now, the argument:

*The Narrow Proportionality Argument Against the ANIMAL RIGHTS THESIS*

- (P1) Defensively killing human Nour to save an animal's life is permissible only if it's narrowly proportionate. [Assumption]
- (P2) Killing human Nour in defense of an animal's life is narrowly proportionate only if human Nour's death wouldn't be morally worse than the animal's death. [From the NARROW PROPORTIONALITY THESIS]
- (P3) An animal's death is itself morally less bad than a human's death. [COMPARATIVE HARM THESIS]
- (C1) Therefore, defensively killing human Nour to save an animal's life isn't morally permissible. [From P1, P2, and P3]

Each of these premises has considerable support in the literature. Denying any of them would therefore be extremely controversial. We could risk denying them, and indeed I will suggest a tentative strategy for rejecting (P1).

To begin, I'll point out that (P1) tacitly assumes that defensively killing someone in a narrowly disproportionate way entails impermissibility. That assumption is problematic. To see what I mean, consider the following example:

*Lucky Youngster*

20-year-old Driver will non-culpably kill 80-year-old Pedestrian unless Pedestrian kills Driver. Both Driver and Pedestrian will live equally good lives and die at 100 years of age.<sup>43</sup>

What does the NARROW PROPORTIONALITY THESIS imply about this case? It implies that if Pedestrian killed Driver, she would impose a *worse* harm on Driver than Driver threatens to Pedestrian. This is because if Driver killed Pedestrian he would deprive her of only 20 years of life, but if Pedestrian killed Driver she would deprive him of *80 years* of life.

Some philosophers think this implication is unproblematic. Kerah Gordon-Solmon, for example, endorses it.<sup>44</sup> She makes clear, however, that the degree of responsibility of Driver matters in determining how much defensive harm he's liable to. She claims that if Driver were threatening a fellow 20-year-old with 80 more equally good years to live, then Driver would be

---

<sup>43</sup> This example is a more described variant of an example given in Jeff McMahan, "The Basis of Moral Liability to Defensive Killing," *Philosophical Issues* 15 (1) (2005): 386-405.

<sup>44</sup> Kerah Gordon-Solmon, "Self-Defense Against Multiple Threats," *Journal of Moral Philosophy* 14 (2): 125-133.

liable to be killed by the 20-year-old pedestrian.<sup>45</sup> She also claims that if Driver *culpably* threatened the 80-year-old Pedestrian, then Driver would be liable to be killed.<sup>46</sup>

I think we have good reason to be suspicious of this position. At first glance, it appears to imply that elderly people are almost never permitted to defend themselves lethally against younger people. But that's not quite right on Gordon-Solmon's view because they *would* be permitted to do so if they were facing off against younger people who *culpably* threatened them. But if that's true, then narrow proportionality isn't a requirement for permissible defense, because the 80-year-old Pedestrian is permitted to kill the 20-year-old culpable Driver *even though* she would thereby impose far more harm on Driver than he poses to her. And, at any rate, it strikes me as implausible to think that the non-culpable Driver in *Lucky Youngster* would avoid liability on narrow proportionality grounds if, say, he non-culpably threatened *four* 80-year-old pedestrians, each of whom would live 10 more years of equal quality to Driver.<sup>47</sup>

Perhaps a threat's degree of responsibility is *internal* to narrow proportionality: culpable threats are liable to more harm, non-culpable but responsible threats are liable to less, and so on. In other words, what counts as narrowly proportionate harm in a given case is a combination of the harm you pose *and* how you pose it. Indeed, this is a common and plausible view of narrow proportionality.<sup>48</sup> But if that's true, then humans who culpably threaten animals will be liable to be killed even if killing them would impose more harm on them than they pose to animals. And humans who non-culpably threaten animals (as many factory farmers do) will be liable to be

---

<sup>45</sup> Gordon-Solmon, "Self-Defense Against Multiple Threats," at 127-8.

<sup>46</sup> Gordon-Solmon, "Self-Defense Against Multiple Threats," at 127.

<sup>47</sup> For those keeping track of the math, that's a combined 40 years for the elderly pedestrians and another 80 years for Driver. Since imposing 80 years of lost life is twice as bad as imposing 40 years of lost life, preventing the latter by causing the former appears to violate narrow proportionality.

<sup>48</sup> Gordon-Solmon, "Self-Defense Against Multiple Threats," at 127, fn. 8, suggests that responsibility is internal to, or necessary for, narrow proportionality. I'm assuming the same is true of *degrees* of responsibility with respect to what counts as narrowly proportionate (or disproportionate) in a given case.

seriously injured. Either way, animal rights militancy appears to evade the force of narrow proportionality, and so (P1) is dubious.

Fortunately, we needn't deny any of the premises, including (P1), to salvage the ANIMAL RIGHTS THESIS. Let's assume that the argument above is successful in that it shows why it's impermissible to kill Nour to save an animal's life. Does this show that killing Nour in *Night Shift Zookeeper* is impermissible?

I don't think it does. One reason is that it *misidentifies* what's morally objectionable about killing Nour in *Night Shift Zookeeper*. Consider a similar example:

*Sickly Zookeeper*

Same as *Night Shift Zookeeper*, but this time Nour is sick and has only a week to live. The pigs, by contrast, have decades to live.

Killing Nour in this case wouldn't be narrowly disproportionate. She has fewer long-term interests than the pigs do: less to live for, and less time to live for it. Making this change to Nour doesn't make killing her any more morally seemly, however. (If anything, it makes it seem worse.) This gives us reason to think narrow proportionality isn't what makes it wrong to kill Nour.

There's a more decisive reason to believe this, however. The conclusion of the *Narrow Proportionality Argument* is that "killing human Nour to save *an* animal's life isn't morally permissible." In other words, the argument entails the

CONDORCET THESIS

It's narrowly disproportionate to kill one human to save the life of one animal.

But this says nothing about killing one human to save *multiple* animals, and it's multiple animals at risk of imminent harm in *Night Shift Zookeeper*. The same is true in most cases of animal rights militancy and rescue, to boot. What's more, the CONDORCET THESIS is a claim only about killing. It makes no claims about nonlethal forms of harming. If we considered a variant in which Nour caused the pigs to suffer immensely and we could prevent this by shooting Nour in the leg (a painful injury), it would be difficult to maintain that shooting her in the leg was narrowly disproportionate.<sup>49</sup> Because so many threats to animals are threats of suffering, as is often the case in the enormous industry of factory farming, the narrow proportionality is bound to fail. We should reject it.<sup>50</sup>

### **3.2. Appeal to Diminished Liability**

Maintaining the background theses and assumptions, it is apparent that any account of Nour's case cannot entail that Nour acts *permissibly*, since Nour unjustly goes against the rights of the pigs.<sup>51</sup> What remains possible is that Nour acts *excusably*. To act excusably is, in this sense, to avoid culpability for one's actions. While non-culpable ignorance is not the only means by which Nour can avoid culpability, it is one way. For Nour's case, we might suppose that Nour is unaware that

---

<sup>49</sup> Robert Nozick claims it's impermissible to inflict slight discomfort on a human in order to spare 10,000 animals from extreme suffering. Nozick makes this claim as an example of what we're forbidden to do to non-labile parties. He makes no claim about what we may do to liable parties, such as Nour. See Nozick, *Anarchy, State, and Utopia* (NY: Basic Books, 2013), at 41.

<sup>50</sup> There may be other ways of denying or weakening the ANIMAL RIGHTS THESIS. I lack the space to explore them here.

<sup>51</sup> Any plausible moral theory on which the ANIMAL RIGHTS THESIS is true will condemn Nour's actions, since her actions would be morally comparable to someone guarding a prison filled with humans who will shortly be unjustly killed.

animals have rights, or perhaps she rejects that view for principled reasons that are not obviously implausible.

Defenders of the ANIMAL RIGHTS THESIS might use this in one of two ways. First, to explain why it's permissible to engage in harmful defensive action against Nour, but *less permissible* than it would be against, say, the prison guards. Second, to explain why it's *impermissible* to engage in harmful defensive action against Nour. On the former view, Nour's non-culpable ignorance is not sufficient to eliminate her liability to defensive harm, but it is enough to make her less liable than the guards. On the latter view, Nour's non-culpable ignorance is sufficient to make her not liable to defensive harm.<sup>52</sup> Consider the following case from Gideon Rosen:

...consider an ordinary Hittite lord. He buys and sells human beings, forces labor without compensation, and separates families to suit his purposes. Needless to say, what he does is wrong. The landlord is not entitled to do these things. But of course he thinks he is. Moreover, we may imagine that if he had thought otherwise, he would have acted differently. In that case he acts from moral ignorance in our sense.<sup>53</sup>

As Rosen would have it, the Hittite lord is non-culpably ignorant of the slave's rights because "chattel slavery was *simply take for granted*," and "it would have taken a moral genius to see through to the wrongness of chattel slavery."<sup>54</sup>

---

<sup>52</sup> It's thus an assumption of the diminished liability approach that it's impermissible to defend human slaves *even if* pacifism is false. As I argue below, this is false. Slaves are permitted to defend themselves with violence, if anyone is.

<sup>53</sup> Gideon Rosen. (2002) "Culpability and Ignorance." *Proceedings of the Aristotelian Society* 103 (1): 61-84, pp. 64-65.

<sup>54</sup> Rosen, "Culpability and Ignorance," pp.65 and 66.

I shall assume for argument's sake that Rosen is right about the Hittite lord: namely, that the lord is non-culpably ignorant for his wrongdoing. Still, non-culpable ignorance cannot diminish liability in the slave case or in the guards case, and for the same reason: because slaves and individuals in concentration camps still possess a right to engage in defensive harm against their brutal oppressors, which implies a third-party permission to defensively assist them against their oppressors.<sup>55</sup> As Hadley explains,

...if people who harm animals are not liable in terms of being responsible for unjustified harms without an acceptable excuse, then moral agents who buy and sell, confine, mutilate without anesthetic, infect with disease, kill for pleasure, and otherwise use rights-bearers as tools will not be legitimate targets for proportionate third-party defensive violence. Such a conclusion would be radically at odds with common sense, if the rights-bearers concerned were human animals of comparable capacities, fetuses aside, and makes the claim that animals have valuable lives worthy of protection ring hollow.<sup>56</sup>

Or consider a case in which my kidnapper is a convinced solipsist, though she grants that she treats me in ways that would be morally monstrous if I were real. She tells me that if she were convinced to abandon her solipsism (which she holds quite sincerely), she would immediately set me free and turn herself in. There again, it is implausible to suppose that I lack defensive rights against my kidnapper, and therefore also implausible to suppose that third parties lack a permission to

---

<sup>55</sup> Whitley Kaufman develops a similar case against the view that it's impermissible to harm innocent aggressors in self-defense. See Kaufman (2010). "Self-Defense, Innocent Aggressors, and the Duty of Martyrdom." *Pacific Philosophical Quarterly* 91: 78-96. See also Whitley Kaufman. (2009) *Justified Killing: The Paradox of Self-Defense*. Lanham, MD: Lexington Books, pp.78-79.

<sup>56</sup> John Hadley, "Animals Rights and Self-Defense Theory," p.169.

defensively assist me against my kidnapper. Commenting on the appeal to diminished liability, John Hadley comments:

A related objection is that it is wrong to teach people about animal rights because inculcation of such views may render them liable to violence on behalf of animals.<sup>57</sup>

Imagine making this move with respect to human rights: You can avoid making people liable to defensive harm by keeping them ignorant about human rights. Such concealment is itself morally wrong. Moreover, the objection assumes that by *concealing* the truth of animal rights from human persons, you thereby make or keep them *non-liable* to defensive harm, including third-party defensive harm. If that were true, one could easily make one's children morally immune to defensive harm by maintaining their ignorance about everyone's rights. Then, no matter whom they unjustly attack, they are not liable to defensive harm. This seems implausible. From these considerations, therefore, we should conclude that Nour's non-culpable ignorance is insufficient to make her non-liable to defensive harm or to defensive assistance.

There is another objection in the conceptual neighborhood. As we have seen, what matters is not liability per se, but permissibility. There might be reasons beyond liability that make animal rights militancy impermissible. One such reason is contractarian in nature. Mark Rowlands offers a contractarian defense of animal rights.<sup>58</sup> He also rejects animal rights militancy for contractarian reasons, arguing that militancy is impermissible because it's unlikely to succeed:

---

<sup>57</sup> John Hadley. (2009) "Moral Responsibility for Harming Animals," *Think* 22 (8): 51-55, p.55.

<sup>58</sup> Mark Rowlands. (1997) "Contractarianism and Animal Rights." *Journal of Applied Philosophy* 14 (3): 235-247.

In the impartial position, then, it would be irrational to choose a world that contains, in addition to non-violent, illegal, attempts to change society, animal rights terrorism. Such terrorism is not necessary for achieving the goals that provide its *raison d'être*. If you turn out to be an animal, you will, in all probability, not be helped by terrorist action. And if you turn out to be human, you may be harmed by it. And, any suffering that you thereby endure, is gratuitous and unnecessary. Therefore, in the impartial position, it would be irrational to choose a world that contains animal rights terrorism. Therefore, in the real world, it is immoral to endorse such terrorism. And that, I think, is why animal rights terrorism is not morally acceptable.<sup>59</sup>

Rowlands doesn't deny that the intended victims of those militant acts are liable to defensive harm. All he denies is that it's permissible to defensively harm them. Christopher Tollefsen offers a similar objection to a militancy objection to the view that fetuses have rights, writing:

If this is correct, then the most plausible interpretation, according to common morality, of violence which is directed against abortion providers and clinics, is that such violence is rebellion, for it is political in nature: it is violence in deliberate contravention to the laws of the state, and with reformatory purposes in mind. In its purposes, and in its means, such action is revolutionary. ... It is a requirement upon any just rebellion that [it will likely succeed]. ... Any rebellion against the established laws of a society, carried out on the grounds that they are grossly unjust, which does not involve a force of rebels substantial enough to have hope of success, must degenerate, I believe, into terrorism or guerilla

---

<sup>59</sup> Mark Rowlands. (2002) *Animals Like Us*. London: Verso Press, pp.193.

warfare, both of which are illegitimate considered as rebellion because, *inter alia*, they are uses of violence which have no hope of success.<sup>60</sup>

A likely chance of success is thus a requirement in the view of both Rowlands and Tollefsen. For both of them, animal rights militancy is unlikely to succeed, making the violence inflicted gratuitous and therefore impermissible. Christopher Kaczor makes a similar but somewhat different point along these lines:

...since abortion is legal in most countries, killing an abortionist is an act of vigilante justice which, like other acts of vigilante justice, is *prima facie* wrong... it is tremendously difficult to overcome the presumption against such violence, due to the likelihood of even worse injustices taking place in the breakdown of social order brought about by vigilante justice.<sup>61</sup>

Unlike Rowlands and Tollefsen, Kaczor appeals not to the low likelihood of successfully protecting the victim's rights, but to the high likelihood of violating others' rights. Both objections appeal to underlying moral principles regulating the permissibility of third-party defensive violence, which I shall call the SUCCESS PRINCIPLE and the VIGILANTE PRINCIPLE, respectively. I shall understand these principles as follows:

SUCCESS PRINCIPLE:

---

<sup>60</sup> Tollefsen, "Donagan, Abortion, and Civil Rebellion," p.310.

<sup>61</sup> Kaczor, *The Ethics of Abortion*, p.220.

Third-party defensive violence is permissible against an unjust aggressor only if that violence stands a reasonable chance of success.

VIGILANTE PRINCIPLE:

Third-party defensive violence is permissible against an unjust aggressor only if that violence would not likely bring about worse injustices.<sup>62</sup>

Perhaps the most significant issue with these principles is that they *don't* imply that it's impermissible to kill Nour. In the *Night Shift Zookeeper* case, killing Nour *does* stand a reasonable chance of success and it *isn't* likely to bring about worse injustices.<sup>63</sup> While *Night Shift Zookeeper* is fictional, it's nonetheless a possibly true case in which it would be impermissible to kill her, and thus any true moral view must explain the wrongness. Since the animal rights militant would have a reasonable chance of success and would not likely bring about worse injustices, neither the SUCCESS PRINCIPLE nor the VIGILANTE PRINCIPLE can plausibly explain the impermissibility of killing Nour.

There are further reasons to reject the SUCCESS PRINCIPLE and the VIGILANTE PRINCIPLE. On the SUCCESS PRINCIPLE, recall Rowlands' claim that because there's no reasonable chance of success to stop unjust aggressors who target animals, the violence inflicted upon those unjust aggressors would therefore be gratuitous. But some, like Suzanne Uniacke, deny this claim:

---

<sup>62</sup> A stronger variation of the VIGILANTE PRINCIPLE might be that such violence is permissible only if it would be *unlikely* to bring about worse injustices. On this interpretation of the moral requirement, a mere lack of a positive likelihood is insufficient. What's needed is a *negative* likelihood (i.e., an unlikelihood). Thus, in cases where the probability of bringing about greater injustices hovers at 0.5, or where there is no better reason to believe that greater injustices will be brought about than that they won't be, it's wrong (on the stronger VIGILANTE PRINCIPLE) for third parties to engage in defensive violence.

<sup>63</sup> Or, per the above footnote, it's *unlikely* to bring about worse injustices.

Although harm inflicted in self-defense is not needless simply in virtue of its being unlikely to succeed, as an instrumental aspect of the justification of self-defense the condition of necessary force does require *some* prospect of success. In practice this means that *B* must believe both that the force he uses is capable of fending off *A*'s threat to some extent and also that the degree of force he uses is not excessive. *B* can believe the former even if he thinks the prospect of success is low. (*B* can consistently believe: It is unlikely that I can fend off *z* unless I use force *x*; and there is only a small chance that if I use force *x* I will fend off *z*.)<sup>64</sup>

Is Uniacke right about this? It seems to me that she isn't, but for reasons that help point to an important objection. If *B*'s employment of force *x* is insufficient (either by itself or when coupled with the employment of other defensive measures like *y*) to fend off threat *z* and *B* knows that, then I don't see how *B*'s employment of *x* can count as a *defensive* measure against *z*. That is, if *B* knows that *x* won't help him rebuff the threat, then it seems that *B* isn't using *x* to defend himself against *z*.<sup>65</sup> As a result, *B*'s use of *x* to rebuff *z* cannot be justified as a defense against *z*.

Daniel Statman considers cases where rape victims are incapable of preventing their rapes but still inflict harms on their rapists. He comments that the SUCCESS PRINCIPLE seems to imply that the rape victims have acted impermissibly because they will not succeed. This implication, thinks Statman, is false:

---

<sup>64</sup> Suzanne Uniacke. (2014) "Self-Defense, Just War, and a Reasonable Prospect of Success," in *How We Fight*, ed. Helen Frowe and Gerald Lang. New York: Oxford University Press: 62-74, p.66.

<sup>65</sup> Might *B* *mistakenly* believe that employing *x* will prevent *z*? No, because (per Uniacke's stipulation) *B* knows that *x* won't prevent *z*. Imagine that *B* used something else he believed wouldn't help him successfully defend against *z*, like singing opera. If *B* sang opera, would he be employing it as a defensive measure? The answer, it seems to me, is that he obviously wouldn't be.

The suggestion that, in all these cases, if the victims estimated that their use of force would not prevent the evil, they would be morally prohibited from using force, is unacceptable.<sup>66</sup>

Helen Frowe joins Statman in thinking that the SUCCESS PRINCIPLE would be false if it had these implications:

Requiring capitulation [to unjust aggressors in cases where the harm will not be averted] does indeed seem to be, as Statman claims, a *reductio* of a moral theory of self-defense. But I don't think any plausible account of self-defense—including those based upon the accounts of liability to defensive harm that are under discussion here—would prohibit these killings.<sup>67</sup>

John Hadley considers an example in which Nazi scientists perform hypothermia experiments on Jews, claiming that the Jews were permitted to engage in defensive violence against the Nazis even if they would have failed to avoid the harms of hypothermia.<sup>68</sup>

Statman suggests that we explain the permissibility of defensive harming in these cases by appealing to honor:

...since whenever Aggressor threatens Victim, he is thereby also threatening, secondarily, Victim's honor, Victim is permitted to carry out otherwise immoral measures against

---

<sup>66</sup> Daniel Statman. (2008) "On the Success Condition for Legitimate Self-Defense." *Ethics* 118 (4): 659-686, p.666.

<sup>67</sup> Frowe, *Defensive Killing*, p.111.

<sup>68</sup> Hadley, "Animal Rights and Self-Defense Theory," pp.174-175.

Aggressor even if, by so doing, Victim will not block the primary threat, provided that in this way Victim can reasonably hope to save her honor.<sup>69</sup>

It would take us too far afield to discuss the merits of Statman's proposal. What it shows, however, is that defensive measures that do not avert the primary threat might nevertheless avert some secondary threat and might be rendered permissible on those grounds. Moreover, the SUCCESS PRINCIPLE falsely implies that rape victims, Jews subjected to hypothermia experiments, and slaves subjected to unspeakable torture aren't permitted to defend themselves against the primary threats they face if they stand no reasonable chance of doing so. As Statman and others suggest, we should reject such a view. By implication, we should reject the view that third-party defense against unjust aggressors who target animals is impermissible unless it will reasonably succeed.

To avoid this implication, Rowlands distinguishes between *efforts to change society*, on the one hand, and *acts of rescue*, on the other. He then offers the following example and commentary:

Suppose you live in a society that condones slavery. You, however, having understood the role of the principles of equality and desert in the moral thinking of your society, oppose this. One day you witness a slave owner beating one of his slaves, and you decide to rescue the slave. This sort of situation has a certain, what we can call, *immediacy*: action is called for as a matter of urgency. You can try an alternative approach—for example, seeking to convince your fellow citizens of the injustice of slavery—but this would take time, and will

---

<sup>69</sup> Statman, "On the Success Condition for Legitimate Self-Defense," p.670.

in no way improve the predicament of the slave who is now being beaten. Therefore, you decide to step in and rescue the slave. This is, then, an act of rescue.<sup>70</sup>

These are precisely the conditions of the *Night Shift Zookeeper* case in which Nour must be killed to save the animals. Thus, they plausibly permit killing Nour in that case. Moreover, the circumstances of the slave are transparently analogous to the circumstances of many animals who won't be saved unless a third party intervenes both *now* and *violently*. As Hadley observes,

It is reasonable to suggest that third-party defensive violence on behalf is likely to be effective in many instances, at least on too many occasions for [defenders of the ANIMAL RIGHTS THESIS] to successfully avoid an absurd conclusion.<sup>71</sup>

The Militancy Objection, then, does not appear to be solved by the SUCCESS PRINCIPLE because the principle is either false, too rarely applicable, or implies a permission to kill Nour grounded in preventing a non-primary wrong.

This brings us to the VIGILANTE PRINCIPLE according to which third-party defensive violence is impermissible if it would likely lead to even greater injustices. I argued above that Nour's case is not like this, and thus the VIGILANTE PRINCIPLE doesn't offer grounds to think killing Nour is impermissible. Moreover, the VIGILANTE PRINCIPLE suffers from some of the same problems that afflict the SUCCESS PRINCIPLE. For example, it will doubtfully apply very often, since many acts of animal rights militancy will not lead to greater injustices. However, it seems to me

---

<sup>70</sup> Rowlands, *Animals Like Us*, pp.185-186.

<sup>71</sup> Hadley, "Animal Rights and Self-Defense Theory," p.175.

that the problem with the VIGILANTE PRINCIPLE is that it makes matters *worse* for defenders of the ANIMAL RIGHTS THESIS. To see why, let's begin by considering the following example:

### *American Civil War*

The United States, divided roughly into the Northern States and the Southern States, participate in the many evils of slavery: Slaves are routinely tortured, raped, and killed in ways that violate their basic moral rights. The Northern States object and end slavery in their territory, but the Southern States persist. After many protests and extended negotiations, the Northern States wage a civil war against the Southern States to prevent the evils of slavery. Over the course of the war, many innocent people die, but not as many as are saved.

I shall assume for the sake of argument that the actions of abolitionist vigilantes, such as John Brown when he and freed slaves violently attacked slaveholders at Harper's Ferry, were impermissible.<sup>72</sup> What seems uncontroversial is that Brown's actions would have been permissible had he banded together with the Northern States' army as a combatant against the Southern States, since it was permissible for the Northern States to wage war against the Southern States as doing so was necessary to end the horrors of slavery. With that in mind, consider the following example:

### *Canadian War*

When it comes to harming animals, the people of Canada are having none of it. They therefore end all animal agriculture, factory farms, and the like in the country and create a

---

<sup>72</sup> This implication, too, strikes me as counterintuitive. John Brown fought in defense of slaves, acting more or less precisely as Rowlands envisions. It seems to me that if pacifism is false, then Brown surely acted permissibly.

safe and secure sanctuary for animals in Nunavut. The United States, meanwhile, continues to violate the basic moral rights of animals. After many protests and extended negotiations, Canada declares war on the United States in an effort to prevent the evils of animal abuse. Over the course of the war, many innocent humans die, but more innocent animals are saved.

Animal abuses in the United States are no less frequent than slave abuses were (in fact, they are *more* frequent) and, on the ANIMAL RIGHTS THESIS, not less impermissible with respect to the basic moral rights that are violated.<sup>73</sup>

The VIGILANTE PRINCIPLE doesn't imply that Canada is permitted to wage war against the United States to defend animal rights. However, that's precisely its problem. The principle condemns the actions of vigilante groups like the Animal Rights Militia for killing Nour, but not Canadian Mounties for doing the same thing. It therefore condemns animal rights vigilantism, but not animal rights *warfare*.<sup>74</sup> What's counterintuitive about the ANIMAL RIGHTS THESIS isn't that it permits you to kill Nour as a vigilante, but that it permits you to kill Nour at all. This gives us strong reason to think that the VIGILANTE PRINCIPLE isn't a satisfactory reply to the Militancy Objection.

### **3.3. Endorse Militancy**

---

<sup>73</sup> A war in defense of mistreated animals might in some ways be easier to wage permissibly than a war in defense of slaves. Concerns about wide proportionality, for example, would be less of a concern, since there were millions of slaves but *billions* of abused animals. This will effectively permit a larger number of non-liable persons to be (unintentionally) killed, since the number of animals who would be saved would be considerably larger.

<sup>74</sup> If the VIGILANTE PRINCIPLE *did* condemn animal rights warfare in *Canadian War*, it would also condemn the war waged by the Northern States in *American Civil War*, which is implausible.

Still another option is to deny (P6), which claims that harming Nour is not permissible. Perhaps it is not so implausible to claim that inflicting defensive harm on Nour is permissible. After all, it would be permissible if Nour were guarding *human* victims, and Nour's non-culpable ignorance is insufficient to rule out liability to defensive harm. Nour violates, or intends to violate, an individual's right, and we are therefore obligated to defend that individual against her wrongful aggression.

But are we obligated? Members of the Animal Liberation Front—a group recognized by the U.S. government as a domestic terrorist group—do not kill or even injure animal researchers, but they face severe criminal penalties simply for freeing animals and burning animal laboratories. Such penalties, whether they are exacted or merely risked, would place enormous burdens on individuals. Far greater penalties would be enacted were the individuals in question to *attack* animal researchers, even if it were necessary to save the chimpanzees. A lifetime in prison, execution, and the like are hardly outcomes that are reasonable to expect moral agents to bear, and thus there is no *duty* to engage in defensive assistance in those cases.

Still, *some* moral agents might be obligated to engage in defensive assistance: namely, agents who will not get caught or even run a serious risk of getting caught. Indeed, there are agents for whom there is *no* risk of getting caught (e.g., particularly stealthy animal liberators) or who would face no burdens even if they were (e.g., politically powerful individuals with numerous outstanding favors). In the absence of other burdens, these agents would be obligated to engage in defensive assistance, and therefore obligated to defend the animals against Nour. Moreover, the denial of (P6) does not imply that moral agents are *required* to be militant, only that they are *permitted* to be militant. Cheryl Abbate, for example, seems committed to the view that many of

us are indeed permitted to engage in animal rights militancy against certain humans on the grounds that they are liable to defensive harm:

I conclude that (1) in cases in which human beings are culpable for posing an unjust threat to nonhuman animals, human beings are fully liable to defensive harm and thus are not justified in harming nonhuman animals in order to defend themselves, and (2) in cases in which human beings are morally responsible (but not culpable) for posing an unjust threat to animals, they should at least share in the costs of their actions.<sup>75</sup>

Thus, on Abbate's view, *culpable* threats are *fully* liable to defensive harm, and *non-culpable but morally responsible* threats are *partially* liable. If human beings are liable to such defensive harms, then they are surely liable to them when doing so requires (or is helped by) third-party assistance.

What sort of threat is Nour? She will unjustly kill the pigs unless she is killed.<sup>76</sup> Nour, therefore, is analogous to an unarmed lookout at a human prison who will call armed guards if 'sympathizers' attempt to free the prisoners. Even so, we might inquire whether Nour is a culpable threat. On Abbate's view, what follows if she is a culpable threat?

Since there is fault in both the *act* and the *agent*, [culpable aggressors] are fully liable to be killed and they should suffer the cost of their own wrongdoing by not harming the animal in self-defense, since they could have avoided these costs by refusing to perform the unjust

---

<sup>75</sup> Abbate, "The Search for Liability in the Defensive Killing of Nonhuman Animals," p.108.

<sup>76</sup> Perhaps we should distinguish between *posing* a threat and *aiding* a threat. But even if we do, aiding a threat is surely going to entail liability to defensive harm, at least when one aids a threat in a morally responsible way, as Nour does. For example, perhaps only Assassin threatens to kill you with a gun, but I hold you in place for the kill shot. Moreover, perhaps I am unaware that Assassin is acting wrongly in doing so—for example, perhaps I mistake Assassin for a police officer doing their duty.

action, while, on the other hand, the nonhuman animal did not have a choice. Furthermore, it is a *matter of justice* that these men suffer the costs of their wrongful actions rather than allowing these costs to be imposed upon the nonhuman animals.<sup>77</sup>

But *is* Nour a culpable threat? I suspect so, but I admit to being unsure about this because I am unsure whether Nour's failure to recognize the basic moral rights of animals is an instance of culpable ignorance.<sup>78</sup> Suppose, then, that she isn't culpable. The non-culpable responsibility for posing a threat, on Abbate's view, entails at least *some* liability on Nour's part. It's not clear exactly what Nour is liable to in Abbate's view, but her view seems to imply that it's less than culpable threats. But is it so much less that Nour isn't liable to be *killed*? Plausibly not, for we might suppose that the individual who unjustly kills human prisoners isn't culpable for the threat he poses. Perhaps, for example, he is given compelling but misleading evidence that the human prisoners are liable. Even so, he seems liable to defensive harm. At the very least, it seems permissible to harm or kill him if doing so is necessary to free the innocent prisoners, regardless of whether the executioner is liable. Liability to defensive harm and the permissibility of defensive harm are logically distinct: If Nour is liable, it doesn't follow that it's permissible to harm her; and if it's

---

<sup>77</sup> Abbate, "The Search for Liability in the Defensive Killing of Nonhuman Animals," p.122.

<sup>78</sup> See, for example, section 3.2 where the solipsist and the person uneducated about human rights provide examples of appeals to ignorance. It seems to me that *both* persons are culpably ignorant because two reasonable prima facie reasonable expectations are that individuals recognize that there are sentient individuals distinct from themselves and that it's impermissible to harm such individuals without a sufficiently good reason. It also seems to me that neo-Cartesian views of animals are false and obviously so, or at the very least that it's reasonable to expect most humans to recognize animal sentience since we think it's reasonable to hold most humans morally and legally responsible for animal abuse. However, I can't defend these positions at length here. These disputes are somewhat peripheral, however, since it seems permissible to harm the convinced solipsist whether or not they are culpably ignorant. The same, therefore, will be true of the lookout, Nour, certain animal researchers, and so on.

permissible to harm Nour, it doesn't follow that she's liable.<sup>79</sup> What matters here, on the militancy view, is that since the prison executioner of humans can be permissibly killed, so can Nour.

What is the scope of animal rights militancy? That is, how many individuals are liable to defensive harm? It's hard to say, but it includes workers in other animal-related professions (e.g., zookeepers, animal researchers), workers in other countries, and other animal abusers (e.g., people who abuse their animal companions). However one does the math, the number of real-world individuals who can be permissibly killed under the terrorist view is simply enormous. To make matters worse, that is only a small sampling of cases in the actual world. Militancy has more problems than that. In addition to all these cases, there is an infinite host of possible worlds brimming over with other cases.

Let's be clear on what this means: Agents are permitted to defensively assist animals against many animal researchers, hunters, farmers, (some) combatants in warfare, and poachers. In short, millions of people. And they are permitted to do this *now*. What matters here is not the *number* of people liable to be killed, but the *kind* of people who are liable to be killed.<sup>80</sup> What seems implausible about the militancy option is that implies these people are liable.<sup>81</sup>

---

<sup>79</sup> Someone is liable to harm just in case you would not wrong them by harming them. Thus, even if you harm a liable individual and thus don't wrong them, you might in doing so still wrong *someone else*, and thus it might be impermissible to harm the liable individual. Moreover, even if it's permissible to harm someone, it doesn't follow that they're liable to be harmed. For example, it might be permissible to harm them as an unintended side effect of preventing some immense harm. For more on distinction between liability and actual permissibility, see Helen Frowe. (2014) *Defensive Killing*. New York: Oxford University Press, p.188.

<sup>80</sup> Hadley's appeal to factors like contributory causation appear to commit him to the view that it is the *number* of liable individuals that generates the counterintuitive nature of the Militancy Objection. Indeed, in his abstract, he refers to the problem as the Multiple Inappropriate Targets Problem. Whereas Hadley and I differ in our interpretation of the problematic nature of the Militancy Objection, our interpretations are not incompatible. Moreover, Hadley's concerns appear to overlap with a similar problem in the broader literature on the ethics of self-defense. See Hadley, "Animal Rights and Self-Defense Theory," p.168. For more on the more general problem, see Frowe (2016: 207-209) and McMahan (2011: 24).

<sup>81</sup> Here's an objection: Militancy doesn't permit us to *kill* all of these people (or people in comparable professions), since doing so would in many cases be *disproportionate* to the harms they threaten to impose on animals. For example, some animal researchers neither kill nor physically injure animals. At most, they imprison animals. Killing those animal researchers to free the imprisoned animals, then, would be objectionably disproportionate. But this objection fails for two reasons. First, the objection relies on the dubious assumption that killing (for example) kidnappers is

### 3.4. Endorse Pacifism

What options are left? If we suppose that the ANIMAL RIGHTS THESIS is true, which entails that the Permission Thesis is true, then we must say that animals have certain defensive rights. But if they have certain defensive rights, this raises the question whether they are permitted to defend themselves against Nour. If not, then either Nour did not wrong them, or she did but some other consideration makes it impermissible to defend against her. By hypothesis, Nour would wrong them, which also entails that the animals neither forfeited nor waived their rights. By the Rights Thesis, no utilitarian considerations or other rights trump the animals' rights. As those are the only two possibilities, it's impossible that some other consideration makes it impermissible for the animals to defend against her, which rules out non-culpable ignorance as a basis. Finally, if we suppose that *Nour* is liable to defensive harm, then we must suppose *everyone like Nour* is open to defensive harm, which entails that an immense amount of harming and killing is permissible.<sup>82</sup> If we disabuse ourselves of these options, including the latter option, then we must concede that animals have rights but that they (and therefore we) are not permitted to engage in defensive violence. The final option, then, is that moral agents are required to be pacifists.

On this view, moral agents are forbidden from using violence in all cases, including self-defense and defensive assistance. That may appear to be incompatible with the PERMISSION THESIS,

---

objectionably disproportionate. It isn't at all obvious that *human* abductees aren't permitted to kill their kidnappers if doing so is necessary to escape. Second, the objection assumes that the intuitive implausibility of the militancy view is restricted to its moral implications for killing, but that's false. Suppose that Nour was guarding not animals who will soon be killed but animals who will soon be tortured, and that in order to rescue them it's necessary to torture Nour. The view that Nour is liable to such harm is again counterintuitive.

<sup>82</sup> If pacifism is true, is *anyone* liable to defensive harm? As I pointed out in footnote 59, liability doesn't entail actual permissibility (and nor does actual permissibility entail liability). By implication, *impermissibility* does not entail non-liability. Thus, even if pacifism entails that assault is always impermissible, it doesn't follow that no one is liable to defensive assault.

but it need not be. Instead, it merely shows that there are limits on what agents can permissibly do in defense of others, a conclusion which itself is not altogether surprising. After all, no one supposes that agents are permitted to do just *anything* in other-defense. Thus, pacifism implies that agents are permitted to defensively assist the pigs, but not in violent ways.

Nevertheless, pacifism would impose very restrictive limits on the PERMISSION THESIS, and would also require some potentially radical revisions to commonsense beliefs about defensive violence.<sup>83</sup> Consider the broad host of wrongful aggressors pacifism forbids us to harm or kill in the actual world. According to the FBI, there were approximately 1,165,383 reported violent crimes in the United States in 2014. A little over 1% of those crimes were murders, and a little over 7% were rapes.<sup>84</sup> These numbers, moreover, reflect only *reported* violent crimes in *one year* in *one country*, and still the number of violent offenses totals well over one million.<sup>85</sup> Now, consider that pacifism implies that third parties weren't permitted to defend these victims with violence.

What seems apparent, then, is that pacifism is as controversial a moral thesis as militancy is, but for different reasons. The apparent problem with militancy is that it implies that we are permitted to attack too many individuals, whereas the apparent problem with pacifism is that it implies that we are permitted to attack too few.<sup>86</sup> The central problems with pacifism and militancy

---

<sup>83</sup> Does pacifism *falsify* the Permission Thesis? As stated, the Permission Thesis merely claims that there's a third-party permission to defend animals. Typically, invoking a permission to other-defense is a way of invoking a permission to other-defensive *assault*. On that reading, pacifism falsifies the Permission Thesis. However, since there are substantive ways of defending others that do not involve assault, it seems that pacifism does not rule out every form of other-defense, and thus is compatible with a broader interpretation of the Permission Thesis.

<sup>84</sup> U.S. Department of Justice, Federal Bureau of Investigation. (2014) "Violent Crime," *Crime in the United States*. URL: <https://www.fbi.gov/about-us/cjis/ucr/crime-in-the-u.s/2014/crime-in-the-u.s.-2014/offenses-known-to-law-enforcement/violent-crime>

<sup>85</sup> The number also reflects only crimes that actually *occurred*, and thus excludes attempted but unsuccessful violent crimes.

<sup>86</sup> This criticism is from Jan Narveson, who argues that pacifists "have too many friends" and "terrorists" (which I'm calling "militants") "have too many enemies." On Narveson's view, terrorism is wrong as a conceptual matter, since it condones the targeting of noncombatants who are not liable to defensive (or offensive) harm. See Jan Narveson.

are therefore two sides of the same coin. Again, with pacifism, the central problem is that it implies we are *forbidden* to harm or kill individuals it seems obvious we are *permitted* to harm or kill. With militancy, the central problem is that it implies we are *permitted* to harm or kill individuals it seems obvious we are *forbidden* to harm or kill. The central problem of both views, then, is *discrimination*: Pacifism forbids us from killing would-be murderers in self-defense, but militancy permits us to slit the throat of a 17-year-old working a night shift at the zoo. These problems are central to pacifism and militancy in the sense that they are the most frequently cited and counterintuitive problems of the views.

Fortunately, the apparent epistemic impasse is resolvable. Pacifism and militancy are on an *initial* epistemic par. The former is saddled with the burden that we cannot permissibly harm a great many aggressors when it seems obvious that we can. Pacifism also has the modal implication that no possible wrongful aggressor can be permissibly harmed, even in possible worlds where doing so is the only feasible means of saving oneself or others. Militancy is thought to be false because it entails that we can permissibly harm individuals who seem morally immune from defensive attack. With militancy, too, the modal implications appear damning: worlds with immense numbers of animal abusers may be harmed or killed to protect the animals. We might even divide the actual and merely possible cases into two broad categories: harming human aggressors who will harm other human beings, and harming human aggressors who will harm animals. Militancy permits both, whereas pacifism forbids both. Our commonsense moral beliefs suggest that harming members of the former group is presumptively permissible whereas harming

---

(2003) "Terrorism and Pacifism: Why We Should Condemn Both." *International Journal of Applied Philosophy* 17 (2): 157-172. Again, the worry here is not that the sheer number of persons liable to be defensively harmed is too many, as if there were some numerical threshold above which there is no liability to defensive harm. Rather, the worry is that these people don't seem to be liable to defensive harm at all, and since the militancy view implies otherwise in so many cases, it gets the wrong judgment in an immense number of cases.

members of the latter group is impermissible. Pacifism implies that our first commonsense belief is false while affirming the second, whereas militancy does just the opposite.

The epistemic par, therefore, is apparent. Other things being equal, we have as good a reason to accept militancy as pacifism.<sup>87</sup> Because of this, we should affirm pacifism, and we need not appeal to any especially controversial moral principle to do so. I shall defend two moral principles that can break the epistemic parity between pacifism and militancy. In the end, these principles support pacifism. The first principle is what I shall call the ASSUME ENDURED RIGHTS PRINCIPLE, according to which if an individual had rights prior to now and it's now unclear whether they still have those rights, we ought to presume they do. It can be represented as follows:

ASSUME ENDURED RIGHTS PRINCIPLE:

If I know Threat had rights prior to now the possession of which would, if Threat still had them, make it actually impermissible for me to harm Threat, and if there are now roughly equally good reasons to believe that Threat currently has those rights and that Threat doesn't, then it's impermissible for me to harm Threat.

---

<sup>87</sup> Helen Frowe (in correspondence) objects that the epistemic par isn't apparent. She concedes that there might be epistemic parity 'downstream', at the level of applied cases, but denies that there's epistemic parity 'upstream', at the level of broad principles and ethical theory. Sometimes we use the independent plausibility of cases to determine the plausibility of moral theories and principles, and sometimes we use the independent plausibility of moral theories and principles to guide us through hard cases. Frowe claims we should rely on the independent plausibility of liability, like we did with the *Burning Barn* case in section §2, to guide us through hard cases like *Night Shift Zookeeper*. Once we do, militancy will be the more plausible view. I can't offer an extensive reply here, but I'll offer one reply. Where militancy and pacifism part ways is at the crossroads of liability, and so (unless we're to beg the question against either view) we need to look elsewhere for an independently plausible moral theory or principle that advantages militancy. Where might it be? So far as I can tell, the sort of 'pre-liability' moral story of the militancy view is *the very same moral story* as pacifism: Both accept that Nour and the pigs have rights, that there's a presumptive obligation against transgressing them, and so on. But if that's true, then the epistemic parity appears *thoroughgoing*: There's intuitive parity downstream with liability, and there's parity upstream due to identical background moral assumptions.

To see why the principle is true, consider an example in which it's extremely controversial whether it's permissible to kill someone defensively. Frowe considers a case like this as an objection to her view on defensive liability. She writes:

I deny that there is a causal threshold for liability to defensive killing, such that one must make a significant contribution to an unjust threat if one is to be liable to defensive killing. But some people have objected that this means that taxpayers who finance an unjust war will be liable to attack on my view.<sup>88</sup>

Frowe goes on to defend her view from this objection, but she claims to recognize that her view is very controversial.<sup>89</sup> Let's assume for argument's sake that it's in fact unclear whether taxpayers are liable to defensive harm. What, then, should be our default assumption? Our options seem to be as follows:

*Option A: Assume Liability*

We can't discern whether taxpayers are, in fact, morally liable to defensive harm. However, we will assume that they are and kill them. If we're right, then we've not wronged *them* (though we may have acted wrongly for reasons independent of liability). If we're mistaken, then we've wronged them (but may have acted permissibly for liability-independent reasons).

*Option B: Assume Non-Liability*

---

<sup>88</sup> Frowe, *Defensive Killing*, p.209.

<sup>89</sup> For her reply to the Taxpayer Objection, see Frowe, *Defensive Killing*, pp.209-212.

Our epistemic situation is the same as in Option A. However, we will assume that taxpayers aren't liable to defensive harm and thus avoid killing them. If we're right, then we've avoided wronging them (though, again, we may have still acted wrongly for liability-independent reasons). If we're mistaken, then we've not wronged them by failing to kill them (but may have acted impermissibly for liability-independent reasons).

What ought we to prefer, morally speaking? In cases where the moral status of an individual is in doubt, some argue that we are morally culpable for killing that individual. For example, Alex Guerrero defends what he calls the DON'T KNOW, DON'T KILL PRINCIPLE:

DON'T KNOW, DON'T KILL PRINCIPLE:

If someone knows that she doesn't know whether a living organism has significant moral status or not, it is morally blameworthy for her to kill that organism or to have it killed, unless she believes that there is something of substantial moral significance compelling her to do so.<sup>90</sup>

Guerrero remarks on the principle:

The justification component is somewhat less straightforward. [Don't Know, Don't Kill] is a caution principle; it requires one to exercise caution before performing certain actions when one is in certain epistemic states, unless one has some justification for going ahead

---

<sup>90</sup> Alexander A. Guerrero. (2007) "Don't Know, Don't Kill: Moral Ignorance, Culpability, and Caution." *Philosophical Studies* 136 (1): 59-97, pp.78-79.

and acting anyway. The justification required in DKDK is that the person believes that there is something of substantial moral significance compelling her to act.<sup>91</sup>

The language of “substantial moral significance” is relevant here, however, because it might be seen as a reason for adopting animal rights militancy over animal rights pacifism. When humans unjustly attack animals, something of substantial moral significance will be sacrificed in either case: the life or welfare of either the human or the animal (or both). Thus, it might be thought that this renders Guerrero’s principle moot with respect to the relevant cases in this paper, including *Night Shift Zookeeper*.

I shall assume for argument’s sake that Guerrero’s principle is moot for this reason. However, it seems to me to be moot for another, more important reason: Neither the direct moral status of the human aggressor nor the direct moral status of the animal is in question.<sup>92</sup> What’s in question is not whether they ever had rights, but whether they *still* have them. On the view that animal rights militancy is permissible, it will (typically) be maintained that the unjust human aggressors have lost their right not to be killed. On the view that animal rights pacifism is true, it will (typically) be maintained that unjust human aggressors haven’t lost that right. What’s agreed upon by *both* views is that both the human aggressors and the animals *had* rights prior to the incident in question.

In my view, we ought to prefer Option B over Option A because our background moral knowledge better supports it. In this case, our background moral knowledge includes knowledge of the fact that the human aggressors and the animals had the right not to be killed prior to the incident. The current epistemic parity is such that we lack sufficiently good reason to believe that

---

<sup>91</sup> Guerrero, “Don’t Know, Don’t Kill: Moral Ignorance, Culpability, and Caution,” p.80.

<sup>92</sup> That is, assuming my arguments in section 3.1 are successful.

anyone lost the rights they had. If that's true, then we lack sufficiently good reason to believe that they're liable to defensive harm. Thus, we should assume non-liability instead of liability.<sup>93</sup> The case for animal rights militancy rests on the assumption that unjust human aggressors are liable or at least should be treated as if they are, and thus a true moral principle (like the Assume Endured Rights Principle) implying otherwise entails the falsity and hence impermissibility of animal rights pacifism. Since our options are animal rights militancy or animal rights pacifism, we should endorse animal rights pacifism. There's another reason to go this route. I shall offer a defense of one further reason:

COMPARATIVE THRESHOLD PRINCIPLE:

If the permissibility threshold is higher for  $k$  than for  $a$ ,  $M$  implies the permissibility of  $k$  and  $P$  implies the permissibility of  $a$ , there's no view that's all-things-considered more plausible than  $M$  or  $P$ , and there's otherwise epistemic parity between  $M$  and  $P$ , then  $P$  is more plausibly true than  $M$ .

This principle, like the ASSUME ENDURED RIGHTS PRINCIPLE, is also a precautionary principle. But it's a different principle. To see why, I'll fill in the blanks further.

The permissibility threshold is higher for killing than for not saving.<sup>94</sup> Consider

---

<sup>93</sup> Here's an objection: Our background knowledge *also* tells us that the human aggressor is aggressing *unjustly*, which is reason enough to prefer militancy over pacifism. But the objection fails because knowing that someone acts unjustly isn't sufficient to know they're liable to defensive harm or that it's permissible to kill them. For example, I might know that you acted unjustly by failing to pay someone what you owed them, but I don't thereby know that you're liable to defensive harm or that it's permissible to kill you.

<sup>94</sup> It's unclear to me whether this is logically distinct from the claim that killing is intrinsically worse than letting die. James Rachels appears to think so in his famous "Active and Passive Euthanasia," in *Applied Ethics*, ed. Peter Singer. (New York: Oxford University Press, 1986): 29-35. Rachels compares two examples in which you either drown your cousin or let him drown. He affirms both are impermissible but denies one violation is intrinsically worse than the other. Warren Quinn argues that even if Rachels is right that neither violation is worse than the other, it doesn't follow

### *Drowning Stranger*

Stranger is drowning and will die unless you jump in the water, swim over to Stranger, and pull him back to shore. But the waters are ice-cold and you'll develop a four-month bout of pneumonia if you do this.<sup>95</sup>

In this case, you aren't morally obligated to save Stranger. The costs to you are too great, though admittedly not as great as they are for Stranger. (Loss of life is worse than pneumonia.) Consequentialists like Peter Singer appear to concede that unequal but morally *comparable* costs can justify not saving others.<sup>96</sup> Comparability, analogous to proportionality, is too strong, however.<sup>97</sup> Suppose you and I are both twenty years of age and both of us will live equally good lives until we're one hundred years of age. Suppose you kidnap me and propose to leave me moderately malnourished and unhappy for ten years, after which you'll set me free. Ten years of moderate malnourishment and unhappiness is hardly comparable to losing eighty years of good life, but I'm still permitted to kill you on commonsense moral views of self-defense. Avoiding nothing less than comparable costs, then, isn't necessary to justify not saving someone. It is, however, the standard for permissible killing. To see why, consider

---

that the permissibility thresholds are the same. See Quinn, "Actions, Intentions, and Consequences: The Doctrine of Doing and Allowing," *The Philosophical Review* 98 (3) (1989): 287-312, at 289.

<sup>95</sup> This is a modified example of Peter Singer's classic example in "Famine, Affluence, and Morality," *Philosophy and Public Affairs*, Vol. 1, No. 3 (1972): 229-243, at 231.

<sup>96</sup> Singer, "Famine, Affluence, and Morality," at 231 and 241.

<sup>97</sup> Cf. Victor Tadros, *The Ends of Harm: The Moral Foundations of Criminal Law* (New York: Oxford University Press, 2011), at 252. Judith Jarvis Thomson defends the stronger view that if even your bodily liberty is at stake, such that it's necessary for you to remain plugged into a violinist for nine months to save his life, that's sufficient to justify not saving his life because the cost to you is too great. See her "A Defense of Abortion," *Philosophy and Public Affairs*, Vol. 1, No. 1 (1971): 47-66, at 52.

*Life Preserver*

You've fallen into a pond and will face a four-month bout of pneumonia unless you can make my way to shore. The only way you can do this is if you pull Stranger into the pond and use their body as a floatation device. This will kill Stranger, but save you.

You act impermissibly if you treat Stranger this way. We might pick any number of explanations for this—for example, that Stranger isn't liable, or that it's impermissible to exploitatively kill non-liable parties, or that you cause far greater to another than you prevent for yourself—but it remains abundantly clear that you're morally forbidden from killing Stranger in *Life Preserver*.<sup>98</sup> What's permissible in *Drowning Stranger* and what's permissible in *Life Preserver* are different: You're permitted not to save Stranger in order to avoid a four-month bout of pneumonia, but not permitted to kill him for that same reason. Because the permissibility threshold is met for not saving in *Drowning Stranger* but not met for killing in *Life Preserver*, and because the cases are otherwise the same, it follows that the permissibility threshold is *higher* for killing than for letting die.<sup>99</sup>

Continuing with my defense of the COMPARATIVE THRESHOLD PRINCIPLE, it should be obvious that militancy implies the permissibility of killing whereas pacifism implies the permissibility of letting die. If false, militancy permits the impermissible killing Nour in *Night Shift Zookeeper*. If false, pacifism prohibits the saving of tortured, guarded prisoners, since doing so requires killing the person guarding them. If my arguments in sections 3.1 and 3.2 are right, there's not a more plausible alternative to militancy and pacifism.

---

<sup>98</sup> For similar cases and discussion, see Frowe, *Defensive Killing*, at 51-71.

<sup>99</sup> The distinction has been defended extensively elsewhere. For defenses, see McMahan, *Killing in War*, at 94; F.M. Kamm, *Intricate Ethics* (New York: Oxford University Press, 2007), at 17; and Quinn, "Actions, Intentions, and Consequences: The Doctrine of Doing and Allowing."

Under uncertainty, we don't know whether it's permissible to kill Nour (as militancy claims) or impermissible to save the prisoners (as pacifism claims). But we *do* know that it's *harder* to meet the threshold for permissible killing than for permissible letting die, and so the view on which there's permissible killing is *intrinsically less likely to be true* than the view on which there's permissible letting die. In other words: Because it's intrinsically harder to justify killing than letting die, any view on which there's justified killing is a view on which (comparatively, under uncertainty) successful justification is intrinsically less likely and, thus, intrinsically more likely to be unjustified. The view that's intrinsically more likely to be true (i.e., the view on which the commitments are intrinsically more plausible) is the view we should hold, at least under uncertainty. That's pacifism. Said still another way, we don't know which possible world we're in: The world in which pacifism is true, P, or the world in which militancy is true, M. Assuming epistemic parity, we should believe P and M are equidistant worlds. But given that P's commitments are intrinsically more likely than M's commitments, P is a likelier closer possible world than M, and the closer world is the one we should believe is the actual world.

Let me argue the point somewhat differently. Militancy might turn out to be the right moral view. So might pacifism. Whichever turns out to be true, their judgments will also turn out to be true, including their counterintuitive judgments regarding Nour and the prison guard. So, neither view has an advantage over the other there. But militancy might turn out to be false, and so might pacifism. Whichever turns out to be false, it'll presumably be because of misjudgments of the kind we're considering: Militancy permits killing when it should prohibit it, and pacifism forbids saving when it should permit (or even require) it. But killing, when it's wrong, is worse than not saving. Thus, militancy risks approving *worse* moral wrongs than pacifism does. This matters when we're

choosing which view to believe under uncertainty because the view that permits the greater wrong is the less plausible view.

#### **4. Conclusion**

Do animals have robust basic moral rights? If they do, those rights include the right to self-defense and the entailed third-party permissions. This appears to imply a presumptive permission to kill any number of humans in defense of vulnerable animals whose basic rights are threatened, a problem I dub the Militancy Objection. I then raised four possible replies and rejected three of them, concluding that the Militancy Objection is best solved by animal rights pacifism.

#### **References**

Abbate, Cheryl. (2015) "The Search for Liability in the Defensive Killing of Nonhuman Animals."

*Social Theory and Practice* 41 (1): 106-130.

Dworkin, Donald. (1984) "Rights as Trumps." Chap. 6 in *Theories of Rights*, ed. J. Waldron. New York: Oxford University Press.

Ebert, Rainer and Tibor R. Machan. (2012) "Innocent Threats and the Moral Problem of Carnivorous Animals." *Journal of Applied Philosophy* 29 (2): 146-159.

Fabre, Cécile. (2006) *Whose Body Is It Anyway? Justice and the Integrity of the Person*. New York: Oxford University Press.

Francione, Gary L. (2000) *Introduction to Animal Rights: Your Child or the Dog?* Philadelphia: Temple University Press.

Frowe, Helen. (2014) *Defensive Killing*. New York: Oxford University Press.

Gardner, Molly. (2016) "The Interspecies Killing Problem." In *The Moral Rights of Animals*, eds. Mylan Engel Jr. and Gary Lynn Comstock (Lanham, MD: Lexington, 2016), 119-140.

Gordon-Solmon, Kerah. (2017) "Self-Defense Against Multiple Threats," *Journal of Moral Philosophy* 14 (2): 125-133.

Guerrero, Alexander A. (2007) "Don't Know, Don't Kill: Moral Ignorance, Culpability, and Caution." *Philosophical Studies* 136 (1): 59-97.

Gunasekera, Crystal Allen. (2018) "The Ethics of Killing 'Surplus' Zoo Animals," *Journal of Animal Ethics* 8 (1): 93-102.

Hadley, John. (2009) "Animal Rights and Self-Defense Theory." *Journal of Value Inquiry* 43: 165-177.

Hadley, John. (2009) "Moral Responsibility for Harming Animals." *Think* 22 (8): 51-55.

Hadley, John. (2009) "Animal Rights Extremism and the Terrorism Question." *Journal of Social Philosophy* 40 (3): 363-378.

Hosein, Adam. (2014) "Are Justified Aggressors a Threat to the Rights Theory of Self-Defense?" In *How We Fight*, ed. Helen Frowe and Gerald Lang. New York: Oxford University Press, 87-103.

Kaczor, Christopher. (2015) *The Ethics of Abortion: Women's Rights, Human Life, and the Question of Justice (Second Edition)*. New York: Routledge.

Kamm, F.M. (2007) *Intricate Ethics*. New York: Oxford University Press.

Kaufman, Whitley. (2010). "Self-Defense, Innocent Aggressors, and the Duty of Martyrdom." *Pacific Philosophical Quarterly* 91: 78-96.

Kaufman, Whitley. (2009) *Justified Killing: The Paradox of Self-Defense*. Lanham, MD: Lexington Books.

Korsgaard, Christine. (1986) "The Right to Lie: Kant on Dealing with Evil." *Philosophy and Public Affairs* 15 (4): 325-349.

McMahan, Jeff. (2016) "The Comparative Badness for Animals of Suffering and Death." In *The Ethics of Killing Animals*, ed. Tatjana Višak and Robert Garner. New York: Oxford University Press, 65-85.

McMahan, Jeff. (2011) *Killing in War*. New York: Oxford University Press.

McMahan, Jeff. (2005) "The Basis of Moral Liability to Defensive Killing," *Philosophical Issues* 15 (1): 386-405.

McMahan, Jeff. (2002) *The Ethics of Killing: Problems at the Margins of Life*. New York: Oxford University Press.

Narveson, Jan. (2003) "Terrorism and Pacifism: Why We Should Condemn Both." *International Journal of Applied Philosophy* 17 (2): 157-172.

Nozick, Robert. (2013) *Anarchy, State, and Utopia*. New York: Basic Books.

Parfit, Derek. (2011) *On What Matters: Volume 1*. New York: Oxford University Press.

Quinn, Warren S. (1989) "Actions, Intentions, and Consequences: The Doctrine of Doing and Allowing." *The Philosophical Review* 98 (3): 287-312.

Rachels, James. (1986) "Active and Passive Euthanasia." In *Applied Ethics*, ed. Peter Singer. New York: Oxford University Press, 29-35.

Regan, Tom. (1983) *The Case for Animal Rights*. Berkeley: University of California Press.

Rodin, David. (2002) *War and Self-Defense*. New York: Oxford University Press.

Rosen, Gideon. (2002) "Culpability and Ignorance." *Proceedings of the Aristotelian Society* 103 (1): 61-84.

Rowlands, Mark. (2002) *Animals Like Us*. London: Verso Press.

Rowlands, Mark. (1997) "Contractarianism and Animal Rights." *Journal of Applied Philosophy* 14 (3): 235-247.

Statman, Daniel. (2008) "On the Success Condition for Legitimate Self-Defense." *Ethics* 118 (4): 659-686.

Tollefsen, Christopher. (1997) "Donagan, Abortion, and Civil Rebellion." *Public Affairs Quarterly* 11 (3): 303-312.

Uniacke, Suzanne. (2014) "Self-Defense, Just War, and a Reasonable Prospect of Success." In *How We Fight*, eds. Helen Frowe and Gerald Lang. New York: Oxford University Press, 62-74.

Uniacke, Suzanne. (1994) *Permissible Killing: The Self-Defense Justification of Homicide*. New York: Cambridge University Press.

## Multiple Threats and the Specter of Pacifism

**ABSTRACT:** The Problem of Multiple Threats (PMT) is a problem for the Standard View of self-defense, according to which there are liability justifications for defensive harming. Against this view is the Pacifist View, which denies the existence of liability justifications. According to the PMT, the Standard View implausibly implies that it's permissible to kill uncountably many liable, threatening individuals. I argue, first, that the strength of the PMT presents a serious and often understated challenge for the Standard View itself. Second, I consider three recent attempts to defuse the PMT in defense of the Standard View and argue that each fails. Third, I show that the PMT is intractable for the Standard View because of what I call the Ratio Principle: It's permissible for Defender to defensively kill unjust threats in approximate proportion to the number of people who would otherwise be unjustly killed by the unjust threats. While the Ratio Principle is non-committal about the precise ratio, it's not less than 1:5, and thus one is permitted to kill at least five unjust threateners to save one's own life. However, this implausibly implies that, in accordance with the Ratio Principle, it's permissible to kill 5 million liable parties to save 1 million non-liable parties. Minimally responsible unjust threateners should have their interests discounted only marginally more than non-liable parties, such that the 1:5 ratio is too strong. Since the PMT is severely and intractably problematic for the Standard View, there's strong reason to prefer the Pacifist View.

### 1. Introduction

Within the ethics of self-defense, the consensus is that the Standard View of self-defense is true. According to the Standard View, there are liability justifications for defensive harming: Some people forfeit their right not to be harmed and, in virtue of that forfeiture, there's a defeasible justification for defensively harming them. According to the chief alternative to the Standard View, the Pacifist View, there are no liability justifications for defensive harming. The consensus is that the Pacifist View is too restrictive in that it doesn't permit *enough* defensive harming.

This paper offers an extended criticism of the Standard View and, with it, a defense of the Pacifist View. I focus on the Problem of Multiple Threats, according to which the Standard View is false because it implies that someone (call them Defender) has liability justifications for killing a limitless number of unjust threats to Defender. If true, then the Standard View permits *excessive* defensive harming.

What precisely is the problem in the Problem of Multiple Threats? Here, it's necessary to offer some background. In this section, I shall accomplish three things. First, I shall introduce the

Problem of Multiple Threats in a general way and explain why it's a problem for the Standard View. Second, I shall state and explain the three leading accounts of moral liability to defensive harm: the RIGHTS ACCOUNT, the CULPABILITY ACCOUNT, and the RESPONSIBILITY ACCOUNT. Third, I shall show how the Problem of Multiple Threats is a problem for each of the views of liability.

Before we begin, however, it's worth introducing a taxonomy of different kinds of threats, since all of the views discussed in section 1.2 revolve around *what kinds* of threats are liable to defensive harm. A *non-responsible threat* is defined as someone (or something) whose threat isn't generated by their moral agency, either because they lack moral agency altogether or because they are typically moral agents but (through no decision of their own) they lack causal control over their currently threatening actions. For example, someone involuntarily pushed from the roof of a house who will crush you, or a rock tumbling from atop a hill, are non-responsible threats. A *minimally responsible threat* is someone whose threat is generated by their moral agency, but without minimal awareness about the threat they pose. For example, an attentive driver who obeys the (reasonable) rules of the road but accidentally threatens a pedestrian crossing the street is a minimally responsible threat. A *(merely) responsible threat* is defined as someone whose threat is generated by their moral agency with full awareness, but without blameworthiness. For example, someone who threatens "as a result of a mistaken belief that doing so is necessary to save their own life" is an example of a responsible threat.<sup>100</sup> A *minimally culpable threat* is someone whose threat is generated by their moral agency with full awareness and with minimal blameworthiness. For example, a driver who speeds slightly through a neighborhood to get to work on time is blameworthy to some extent for doing so. A *fully culpable threat* can be defined as someone whose threat is generated by their moral agency with full awareness and 'robust' or non-minimal

---

<sup>100</sup> Helen Frowe, *Defensive Killing* (New York: Oxford University Press, 2014), at 1.

blameworthiness. For example, a vengeful person intent on destroying his enemy by killing his work rival's innocent children is a fully culpable threat.

### 1.1. Introducing the Problem

In her work on the ethics of defensive harming, Helen Frowe remarks that most people think there are moral exceptions to the moral prohibition against “non-consensually harm[ing] other people.”<sup>101</sup> The paradigm exception is *self-defense*, and those who endorse this exception endorse what I'll call the Standard View. Those who don't think self-defense is an exception endorse the Pacifist View.<sup>102</sup>

The contemporary literature on the ethics of self-defense bears out the sheer dominance of the Standard View. Jeff McMahan, for example, refers to what I call the Standard View as “an axiom of contemporary moral theory,”<sup>103</sup> and Tyler Doggett's 2011 survey article bypasses the debate over Pacifism and focuses exclusively on “*why* you can kill in self-defense” and “*which things* you can kill in which cases.”<sup>104</sup> Importantly, Doggett's review of answers to the latter question doesn't include Pacifism, since it's relatively uncontroversial amongst the primary three views—the Rights Account, the Culpability Account, and the Responsibility Account—that it's permissible to kill *some* individuals in defense of self or others.<sup>105</sup>

---

<sup>101</sup> Frowe, *Defensive Killing*, at 21.

<sup>102</sup> Pacifism applies more broadly than to self-defense, as it's a thesis about non-consensual harming more generally. Nevertheless, the extremely limited view of self-defense is a pacifist one.

<sup>103</sup> Jeff McMahan, “Self-Defense and the Problem of the Innocent Attacker,” *Ethics* 104 (1994): 252-290, at 252. Cf. Kaila Draper, “Defense,” *Philosophical Studies* 145 (1) (2009): 69-88, at 69.

<sup>104</sup> Tyler Doggett, “Recent Work on the Ethics of Self-Defense,” *Philosophy Compass* 6 (4) (2011): 220-233, at 229, my emphases. Notably, then, Doggett thinks the relevant which-thesis and why-thesis won't be null sets.

<sup>105</sup> These views—the Rights Account, the Culpability Account, and the Responsibility Account—are explained below.

The standard justification for self-defense by each of these three accounts is a *liability* justification. To say someone is liable to some (non-consensual) harm is to say they wouldn't be wronged by that harm because they forfeit their right against that harm. Individuals forfeit their right against defensive harms by posing unjust threats, though (as we'll see below) there's much disagreement about what else is needed (e.g., culpability).

This is what's taken to be the principal problem with Pacifism: It fails properly to *differentiate liability-wise* between individuals, and thereby violates an axiom of contemporary moral theory. More particularly, it denies (when it should affirm) that it's all-things-considered permissible to defensively kill certain individuals for reasons grounded in the unjust threats they pose. By sanctioning zero non-consensual defensive killings, it thereby sanctions far too little defensive killing.

Virtually all of the disputes in the contemporary literature on self-defense, then, are disputes *within* the Standard View. As matters turn out, however, the Standard View suffers from a discrimination problem of its own, one sufficiently serious to merit strong reconsideration. Enter one of the cutting-edge moral problems in the ethics of self-defense is what has come to be called the Problem of Multiple Threats. The problem has its origins in the work of Jeff McMahan, who offered the following case:

Suppose, for example, that ten innocent civilians from a country that is fighting a just war have been unjustly imprisoned in the enemy country. It is reasonable to suppose that each military guard at the prison in which they are confined is liable to be killed in the course of a military operation to free

them.... But now suppose that all the guards are known to be reluctant conscripts and that to free the ten prisoners it would be necessary to kill five hundred guards.<sup>106</sup>

McMahan thinks it's impermissible to kill them. But, as he points out, it's not easy to say why since, on his view of defensive liability, even minimally responsible threats can be liable to defensive harms, and it's difficult to see how the mere number of threats could itself make a difference to the permissibility of killing them. This suggests that the Standard View, like Pacifism, has a discrimination problem of its own, but on the opposite side of the coin: namely, that it sanctions *too much* defensive killing. What precisely is the problem in the Problem of Multiple Threats? Here, it's necessary to offer some background.

## **1.2. Three Views of Liability**

As mentioned in the last section, the contemporary ethics of self-defense is dominated by the Standard View, which consists of three primary sub-views: the Rights Account, the Culpability Account, and the Responsibility Account. These accounts serve dual functions. First, they identify *who* can be permissibly defensively harmed. Second, they identify *why* those individuals can be permissibly defensively harmed. Each view attempts to explain why Threat is liable to (i.e., isn't wronged by) some defensive harm. Thus, each view attempts to explain when there's a liability justification for defensively harming Threat. I shall now offer brief characterizations of these accounts. The first account, the Rights Account, maintains the following:

---

<sup>106</sup> Jeff McMahan, *Killing in War* (New York: Oxford University Press, 2009), at 24.

RIGHTS ACCOUNT:

Defender has a liability justification for non-consensually harming Threat if and only if Threat will otherwise violate Defender's right not to be unjustly harmed.<sup>107</sup>

The most long-standing defender of this view is Judith Jarvis Thomson. She begins her classic paper with the case of *Culpable Threat*, in which Villain culpably attacks Defender and is subsequently killed by Defender. Thomson claims that Defender's killing Villain is morally permissible because Villain will otherwise violate Defender's right not to be killed.<sup>108</sup> She extends this explanation to another case, *Non-Responsible Threat*, in which Driver has been faultlessly hypnotized by Villain to kill Defender, claiming again that Defender is permitted to kill Driver.<sup>109</sup>

While she doesn't explicitly consider it, the literature published after Thomson's analysis includes a discussion of a third kind of threat, i.e., the minimally responsible threat. Jeff McMahan has offered a variant of Thomson's driving case:

A person keeps his car well maintained and always drives cautiously and alertly. On one occasion, however, freak circumstances cause the car to go out of control. It has veered in the direction of a pedestrian whom it will kill unless she blows it up by using one of the explosive devices with which pedestrians in philosophical examples are typically equipped.<sup>110</sup>

---

<sup>107</sup> Doggett, "Recent Work on the Ethics of Self-Defense," characterizes the Rights Account differently, i.e., as a forfeiture account. However, most liability-based accounts of defensive harming do this. What's interesting about the Rights Account, as defended by individuals like Thomson, is that they don't restrict permissible killing to culpable or responsible aggressors, but (as we shall soon see) even to *non-responsible* threats. Moreover, my characterization of the Rights Account is present in the literature. See, for example, Jeff McMahan, "The Basis of Moral Liability to Defensive Killing," *Philosophical Issues* 15 (2005): 386-405, at 387.

<sup>108</sup> Judith Jarvis Thomson, "Self-Defense," *Philosophy and Public Affairs* 20 (4) (1991): 283-310, at 283.

<sup>109</sup> Judith Jarvis Thomson, "Self-Defense," at 284. For clarity of exposition, I have opted to change the original names of Thomson's examples. However, they are *Villainous Aggressor* and *Innocent Aggressor*, respectively.

<sup>110</sup> Jeff McMahan, "The Basis of Moral Liability to Defensive Killing," at 393.

Call this case *Single Conscientious Driver*. What's relevant about this case for the Rights Account is that minimally responsible threats would just as surely violate rights as culpable or non-responsible threats. As McMahan puts it: "It does not require culpability, responsibility, or even agency on the part of the violator."<sup>111</sup> Thus, the Rights Account implies the permissibility of defensively harming not only culpable threats, but also merely responsible threats and non-responsible threats. As I show in the next section, this sweeping account of liability makes the Problem of Multiple Threats a severe and intractable problem for the Rights Account.

The next account, the Culpability Account, offers a different view about who's liable to defensive harm:

CULPABILITY ACCOUNT:

Defender has a liability justification for non-consensually harming Threat if and only if Threat is culpable for an unjust threat to Defender.<sup>112</sup>

This account restricts liability justifications to culpable threats, where Threat is a culpable threat if and only if Threat is *blameworthy* for posing an unjust threat. Kimberly Ferzan defends one variant of this view on which Threat is liable if and only if:

(1a) The aggressor forms an intention to purposefully, knowingly, or recklessly kill the defender, the aggressor lacks a justification or excuse, and the defender must kill the aggressor

---

<sup>111</sup> Jeff McMahan, "The Basis of Moral Liability to Defensive Killing," at 389.

<sup>112</sup> Here, I adopt McMahan's characterization of the Culpability Account to include "culpable *responsibility* for an unjust threat," as opposed to the weaker "culpably *posing* an unjust threat." See McMahan, "The Moral Basis of Liability to Defensive Killing," at 390.

to prevent being killed himself, or (1b) the aggressor purposefully, knowingly, or recklessly engages in conduct that he is aware may lead the defender to believe that (1a) is true, and the aggressor lacks a justification or excuse for so doing.<sup>113</sup>

Ferzan's account of when a threat is culpable is a description of when a threat is blameworthy. Despite disagreements about what's needed, blameworthiness is the defining characteristic of a culpable threat. The Culpability Account is therefore the most restrictive account of liability as it sets the bar high for liability, implying that merely responsible and non-responsible threats are never liable.

Finally, there's the Responsibility Account, which presents the following view about liability justifications for defensive harming:

RESPONSIBILITY ACCOUNT:

Defender has a liability justification for non-consensually harming Threat if and only if Threat is morally responsible for an unjust threat to Defender.

Someone can be morally responsible for an unjust threat without being a culpable threat. What distinguishes merely responsible threats from culpable threats is that the former aren't blameworthy whereas the latter are. Merely responsible threats are responsible in the bare agential sense: They act voluntarily (i.e., for their own reasons and without coercion) but not blameworthily.

---

<sup>113</sup> Kimberly Ferzan, "Culpable Aggression: The Basis for Moral Liability to Defensive Killing," *Ohio State Journal of Criminal Law* 9 (2012): 669-697. Cf. Larry Alexander and Kimberly Ferzan, *Crime and Culpability: A Theory of Criminal Law* (New York: Cambridge University Press, 2009), chapter 2.

McMahan's *Single Conscientious Driver* case is an example of this. There, "the driver is liable because he voluntarily engaged in a risk-imposing activity and is responsible for the consequences when the risks he imposed eventuate in harms."<sup>114</sup> This, McMahan contends, is more aligned with considerations of fairness because Driver is minimally responsible for the eventuating harms whereas Pedestrian isn't at all responsible, and so "it is fair that [Driver] should suffer them rather than impose them all on the pedestrian."<sup>115</sup> Thus, on the Responsibility Account, there are liability justifications for killing culpable and minimally responsible unjust threats (since both are morally responsible for posing unjust threats) but no such justification for non-responsible threats (since they aren't responsible, in the agential sense, for posing unjust threats).

Having explained these views, I'll now explain how each of them suffers from the Problem of Multiple Threats. This time, however, I'll take the accounts in reverse, beginning with the Responsibility Account. This is because Problem of Multiple Threats has been aimed primarily at the Responsibility Account.

Why is the Problem of Multiple Threats a problem for the Responsibility Account? The central example is a variant on *Single Conscientious Driver* but involves multiple drivers, and finds its place in a recent and influential essay by Kerah Gordon-Solmon:

Now, 1,000 drivers simultaneously threaten a single (young and healthy) pedestrian. Just in case the pedestrian doesn't kill each driver, her death is overdetermined: it's true of each driver that she will kill the pedestrian unless the pedestrian kills her first. Suppose the pedestrian will in fact kill all the drivers. Each killing will save the pedestrian's life. Because in each case the

---

<sup>114</sup> McMahan, "The Basis of Moral Liability to Defensive Killing," at 394.

<sup>115</sup> McMahan, "The Basis of Moral Liability to Defensive Killing," at 394. Cf. Kerah Gordon-Solmon, "Self-Defense Against Multiple Threats," *Journal of Moral Philosophy* 14 (2017): 125-133, at 128.

driver is minimally responsible and the pedestrian is non-responsible, each killing is narrowly proportionate. So each driver is liable to be killed and each killing is *pro tanto* justified by the driver's liability.<sup>116</sup>

Call this case *Multiple Conscientious Drivers*. The problem, says Gordon-Solmon, "is that it seems wrong to kill 1,000 minimally responsible drivers because the badness of the harms *to the drivers* is excessive."<sup>117</sup> The emphasis on the excessive harm *to the drivers* is, she thinks, a desideratum for a plausible explanation.

Before continuing, it's important to introduce some terminology. First, to say that a harm Defender inflicts on Threat is *narrowly* proportionate (or proportionate in the narrow sense) is to say that the harm doesn't exceed the harm Threat poses. Said another way, "Narrow proportionality is a constraint on individual liability" that "sets the limit to the amount of harm it can be permissible to inflict on an individual on grounds of liability."<sup>118</sup> Thus, if Threat will kill Defender unless Defender kills Threat, then Defender's killing Threat is narrowly proportionate. By contrast, if Threat will kick Defender's leg unless Defender kills Threat, then Defender's killing Threat is narrowly *disproportionate*. Second, to say that a harm Defender inflicts on someone is *widely* proportionate (or proportionate in the wide sense) is to say that the harm isn't "excessive, weighed against the good" it would achieve.<sup>119</sup> Thus, if Threat will kill Defender unless Defender tosses a grenade that will kill Threat and harm no one else, then Defender's tossing the grenade is widely proportionate. By contrast, if the grenade will kill not only Threat but also ten nearby

---

<sup>116</sup> Gordon-Solmon, "Self-Defense Against Multiple Threats," at 128.

<sup>117</sup> Gordon-Solmon, "Self-Defense Against Multiple Threats," at 128-129.

<sup>118</sup> Jeff McMahan, "Liability, Proportionality, and the Number of Aggressors," in Saba Bazargan-Forward and Samuel C. Rickless (eds.) *The Ethics of War* (New York: Oxford University Press, 2017): 3-27, at 4.

<sup>119</sup> Gordon-Solmon, "Self-Defense Against Multiple Threats," at 125.

bystanders, then Defender's tossing the grenade is widely *disproportionate*. Notably, while both widely disproportionate harming and narrowly disproportionate harming count against the all-things-considered permissibility of an action, only the latter counts against a *liability justification* for an action, since whether or not bystanders are harmed doesn't affect *Threat's* liability. Thus, in the grenade case involving the killing of ten nearby bystanders, there is a liability justification for killing only Threat with the grenade, but this justification is defeated by the violation of wide proportionality, making the action all-things-considered impermissible while not mitigating or eliminating Threat's liability.

The problem for the Responsibility Account therefore appears to be that the liability justification for killing all 1,000 drivers is undefeated, since the harms imposed on them are both narrowly proportionate and widely proportionate and, as McMahan claims, this "is implausible in the case of [minimally] responsible killers."<sup>120</sup> This is because, as Gordon-Solmon claims, "The fact that the drivers are not more than minimally responsible is load bearing on this judgment."<sup>121</sup> She calls this moral misjudgment "embarrassing" for the Responsibility Account.<sup>122</sup> Kaila Draper joins the chorus, noting that "it is not at all obvious that it would be proportionate to harm large numbers of aggressors in defense of a single person."<sup>123</sup>

Although the Problem of Multiple Threats has been applied most often to the Responsibility Account, some in the literature claim it also extends to the Culpability Account. As Gordon-Solmon observes, what seems intuitively wrong about the *Multiple Conscientious Drivers* case is that the harm to the drivers seems excessive, relative to their degree of responsibility. She maintains, further, that these harms don't seem excessive "if the threats were fully or partially

---

<sup>120</sup> McMahan, "Liability, Proportionality, and the Number of Aggressors," at 6.

<sup>121</sup> Gordon-Solmon, "Self-Defense Against Multiple Threats," at 132.

<sup>122</sup> Gordon-Solmon, "Self-Defense Against Multiple Threats," at 128.

<sup>123</sup> Kaila Draper, *War and Individual Rights* (New York: Oxford University Press, 2017), at 118-119.

culpable” even if there are multiple drivers.<sup>124</sup> There’s reason to doubt this claim, however. Imagine that an underemployed driver living in a suburban area knows the roads are dangerously icy but also knows that if she doesn’t arrive punctually to her low-wage job, she’ll be fired and lose whatever small financial standing she currently has. She therefore drives her car quickly to work and hits some black ice while turning quickly into work and (unjustly) threatens to kill Pedestrian. Call this case *Single Minimally Culpable Driver*.

Several things seem true of this driver. First, she’s more responsible than the minimally responsible driver in *Single Conscientious Driver*, since for the minimally responsible driver it was improbable (on that driver’s evidence) that his car would pose a danger to any pedestrians, whereas for the minimally culpable driver it wasn’t unlikely (again, on her evidence). Second, she’s responsible to a culpable degree: She knowingly poses a significant risk of severe unjust harm to pedestrians. Third, she’s only minimally culpable since it isn’t her aim to kill Pedestrian and her negligent driving is motivated not by malice or apathetic negligence, but by an understandable urgency to keep her job. Because of this, the underemployed driver is a more sympathetic case: Unlike the conscientious driver, the minimally culpable driver culpably fails to discharge all of her relevant obligations; however, she fails to do so in the face of non-trivial moral burdens. Were there 1,000 such minimally culpable drivers, defensively killing all of them would seem somewhat excessive. This shows us two things. First, it shows us that Gordon-Solmon is wrong when she claims that the deaths of multiple culpable threats, whether partially culpable or fully culpable, aren’t sufficiently bad to make it seem impermissible to defensively kill all of them. Second, it shows us that the Culpability Account doesn’t escape the Problem of Multiple Threats, since it implies the all-things-considered permissibility of defensively killing 1,000 minimally

---

<sup>124</sup> Gordon-Solmon, “Self-Defense Against Multiple Threats,” at 132.

culpable drivers, which seems all-things-considered impermissible relative to their very minimal degree of culpability. Lastly, consider the following case:

Jaywalker needs to cross a street where there's no crosswalk in order to make it to work on time. If he isn't punctual, he'll lose his job to an unreasonable employer. Jaywalker knows the street is often busy and has several blind driveways and cross-streets. Seeing no cars, he decides to cross the street. Jaywalker's co-worker, Driver, is on her way to work and is also reasonably concerned about being terminated if she's late. She realizes there's a risk of whipping around the bend too quickly and she sees Jaywalker crossing the street, but she believes Jaywalker is far enough away that she doesn't need to slow down. Both Jaywalker and Driver suddenly realize the imminent danger: Driver will kill Jaywalker unless Jaywalker blows up her car (and her along with it), and the force of Jaywalker's body will cause damage to Driver's car and minimal damage to Driver herself (e.g., whiplash).

Call this case *Jaywalker*. On the Culpability Account, Jaywalker isn't liable to defensive harm from Driver because Jaywalker responsibly but non-culpably poses a threat to Driver. He's fully responsible because he doesn't use a crosswalk, but he isn't culpable because there's no reasonably accessible crosswalk and he checks and sees no oncoming cars.<sup>125</sup> What's more, the only defensive harm Driver could impose on Jaywalker is killing him, which would be narrowly disproportionate given that Jaywalker threatens only minor damage to Driver and her car. Driver, on the other hand, *is* liable to defensive harm from Jaywalker because Driver culpably poses a threat to Jaywalker.

---

<sup>125</sup> I'm assuming that jaywalking in this way isn't itself morally objectionable. Things would be different if there were a reasonably accessible and functioning crosswalk and Jaywalker declined to use it, or if Jaywalker misused a reasonably accessible and functioning crosswalk (e.g., if Jaywalker decided to walk when the signal indicated he shouldn't), or if Jaywalker didn't check for cars before crossing the street, or if Jaywalker were crossing for morally trivial reasons (e.g., because he enjoys imagining himself a badass).

It's understandable why she didn't slow down, given her beliefs about Jaywalker's distance, but she's still culpable for doing so. What's more, Jaywalker's defensive killing of Driver would be narrowly proportionate, since Driver threatens Jaywalker with (unjust) death. The intuitively right judgment in *Jaywalker* is that it's all-things-considered permissible for Jaywalker to defensively kill Driver, but not the other way around.<sup>126</sup> However, the degree of responsibility between Jaywalker and Driver is fairly minimal: Jaywalker is fully responsible yet just shy of culpability, whereas Driver is minimally culpable and therefore just barely above full responsibility. Thus, in a case where Jaywalker must defensively kill 1,000 minimally culpable drivers, it seems impermissible for Jaywalker to defensively kill *all* of them given Jaywalker's comparable degree of responsibility. In other words, the harms to the 1,000 minimally culpable drivers seems excessive relative both to *their* degree of responsibility and to *Jaywalker's* degree of responsibility. Cases like this show that the Culpability Account can't easily avoid the Problem of Multiple Threats.

Lastly, there's the Rights Account. As we saw previously, this account is liberal with respect to liability justifications. The Rights Account requires only that an unjust threat be posed, and *not* that it be posed by an agent. It therefore holds not only that culpable and merely responsible threats (including minimally responsible threats) are liable to non-consensual defensive harms, but also that non-responsible threats are also liable. It's less restrictive than the Responsibility Account, which holds that agential responsibility is a necessary condition for liability, and far less

---

<sup>126</sup> This intuitive judgment holds even if we assume that Jaywalker is minimally culpable for jaywalking, since even if he were culpable for doing so, killing him would be narrowly disproportionate and therefore not all-things-considered permissible. Moreover, the most plausible variation of the Culpability Account will maintain that *degree* of culpability matters for assigning liability, and that it would be impermissible to kill the less culpable party over the more culpable party. For example, if Black Jaywalker is culpable and Racist Driver intends to ram Black Jaywalker because he's black, any plausible version of the Culpability Account should say that it's all-things-considered permissible for Black Jaywalker to defensively kill Racist Driver and not the other way around.

restrictive than the Culpability Account, which holds that culpability is a necessary condition for liability.

Judith Jarvis Thomson is the principal defender of the Rights Account. What's distinctive about her defense, and the Rights Account itself, is its emphasis on posing an unjust threat as *sufficient* for liability to defensive harm. Thomson is explicit that even if a young child accidentally poses an unjust threat, they remain liable to defensive harm.<sup>127</sup> Keeping this in mind, consider my own variant on the driver cases:

Parent is leaving for work and puts Toddler in the backseat. Unbeknownst to Parent, Toddler crawls into the front seat and accidentally puts the car in drive. The car rolls down the hill and Toddler, while playing, unknowingly steers the car toward Pedestrian, who will be killed unless he blows up the car, destroying it and killing Toddler.

Call this case *Driving Toddler*. The namesake is uncontroversially non-responsible for the threat to Pedestrian. Nevertheless, Toddler threatens Pedestrian's right not to be unjustly harmed.<sup>128</sup> On

---

<sup>127</sup> The moral judgment in *Driving Toddler* closely tracks a similar moral judgment by Thomson, easily the most famous defender of the Rights Account. In her defense of abortion, Thomson offers a case in which a child rapidly expands inside a house, threatening another inhabitant. She remarks, "But it cannot be concluded that you too can do nothing, that you cannot attack it to save your life. However innocent the child may be, you do not have to wait passively while it crushes you to death. Like the driving Toddler, the expanding child is a non-responsible threat. This, however, does nothing, in Thomson's view, to mitigate their liability. See Judith Jarvis Thomson, "A Defense of Abortion," *Philosophy and Public Affairs* 1 (1) (1971): 47-66, at 52.

<sup>128</sup> Things are worse than this. Suppose it's not Toddler at the wheel but Baby in the backseat. The car careens toward Pedestrian and Pedestrian can save his life only by destroying the car, which will kill Baby. (Call this case *Backseat Baby*.) Frowe offers an interesting case in which someone wearing heavy ski boots is, as a result of their own weight plus the ski boots, a lethal threat to someone. She notes (*Defensive Killing*, at 39), "It would be an odd moral theory that treats Falling Person as a threat when it is *her* weight that will kill Victim, but not when it is the weight of the ski boots that will kill him. When the threatening object will kill Victim qua object—when it is the physical mass that threatens Victim's life—being part of that threatening mass is morally relevant. Since Guard contributes to the trolley's mass, he contributes to the threat. If Frowe is right, then it appears the Rights Account is even more counterintuitive because Toddler's directing the car toward Pedestrian isn't a morally salient detail. Baby is a threat just in virtue of being a contingent part of the car,

the Rights Account, therefore, Pedestrian has a liability justification for killing Toddler. Imagine now that 1,000 driving toddlers are hurtling toward Pedestrian, each in their own car, and that Pedestrian is capable of killing them all. If killing minimally responsible threats seems impermissible due to their minimal degree of responsibility, it appears all the more impermissible to kill non-responsible threats, who bear exactly zero responsibility for the threats they pose.<sup>129</sup>

The Rights Account is thus in especially hot water. In general, the lower the responsibility-threshold for liability, the more counterintuitive the results are in multiple threat cases. Recall that all that's necessary for liability, on the Rights Account, is unjustly threatening someone's rights, which one can do even without being responsible in the agential sense. Thus, even non-responsible threats can be liable. Some philosophers find this implication troubling independent of the Problem of Multiple Threats.<sup>130</sup> However, it's clear that the Rights Account runs seriously afoul of the Problem of Multiple Threats, as well.

## 2. Worsening the Problem

In this section, I show how the problem is worse than has been recognized in the contemporary literature, namely, that the Problem of Multiple Threats poses a severe discrimination challenge to

---

such that car-with-a-passenger is a composite threatening object to Pedestrian, and thus Pedestrian is permitted to destroy that object, including the passenger (i.e., Baby).

<sup>129</sup> Elizabeth Anscombe famously decried the wrongness of the bombing of Hiroshima and Nagasaki, in which killing thousands of "innocents" (i.e., non-threatening individuals). She asks, rhetorically, Come now: if you had to choose between boiling one baby and letting some frightful disaster befall a thousand people—or a million people, if a thousand is not enough—what would you do?

If we reversed the judgment—boiling or otherwise killing 1,000 babies to save one pedestrian—this would seem all the more impermissible. Revising things even further so that the babies are somehow part of the threat (e.g., they're present at military bases or in military homes), much of the force of Anscombe's point remains. See Anscombe, "Mr. Truman's Decree," in *The Collected Philosophical Papers of G.E.M. Anscombe: Volume III: Ethics, Religion, and Politics* (Oxford: Wiley-Blackwell, 1981): 62-71, at 65.

<sup>130</sup> See, for example, Michael Otsuka, "Killing the Innocent in Self-Defense," *Philosophy and Public Affairs* 23 (1) (1994): 74-94, at 80.

the Standard View. This places the Standard View in a comparable position to the Pacifist View, which has a similar discrimination problem widely regarded as a near-decisive reason to reject it. Due to important internal commitments of the two views, however, this advantages the Pacifist View.

The Problem of Multiple Threats has been as primarily a problem for the Responsibility Account. This is why those principally concerned with it—such as Frowe, Gordon-Solmon, and McMahan—have increasingly focused their attention on it. Nowhere, however, have any of these authors suggested abandoning the Responsibility Account because of the Problem of Multiple Threats.

The problem, however, is worse than it initially appears. The Responsibility Account is housed under the Standard View, the latter of which has achieved something akin to near-consensus. Its contrasting alternative, Pacifism, despite the occasional full-throated defense, has been universally panned.<sup>131</sup> Indeed, it has been pacifism's commitment to Pacifism that has been widely regarded as decisive reason to reject it. Jan Narveson, for example, argues against pacifism on precisely these grounds, writing,

That's what's wrong with pacifism as a moral stance: at bottom, it *fails to distinguish morally between aggressors and their victims*, and so between the innocent and the guilty.<sup>132</sup>

---

<sup>131</sup> For examples of defenses of Pacifism, see: Cheyney Ryan, "Self-Defense, Pacifism, and the Possibility of Killing," *Ethics* 93 (3) (1983): 508-524; Blake Hereth, "Against Self-Defense," *Social Theory and Practice* 43 (3) (2017): 613-635.

<sup>132</sup> Jan Narveson, "Terrorism and Pacifism: Why We Should Condemn Both," *International Journal of Applied Philosophy* 17 (2) (2003): 157-172, at 159 (original emphasis). Cf. Jan Narveson, "Pacifism: A Philosophical Analysis," *Ethics* 75 (4) (1965): 259-271, at 259.

The problem with Pacifism, therefore, is principally one of *liability differentiation*: It denies that there are any liability justifications for non-consensually harming certain individuals, but doing so it profoundly counterintuitive. Some threats, such as Thomson's Villain, seem clearly liable to defensive harm.<sup>133</sup> And this is why next to no one endorses Pacifism.

That's where the Problem of Multiple Threats gains strength, however. Certain defenders of the Responsibility View, such as McMahan, accept the view that there's a stronger moral presumption against killing than against letting die (or not saving). This moral principle, McMahan thinks, "underlie[s] the moral asymmetry between fighting in an unjust war and not fighting in a just war."<sup>134</sup> This doesn't require that killing is morally worse than letting die; rather, it requires only that there's a stronger moral presumption against killing than against letting die.<sup>135</sup> The latter claim is exceedingly plausible. Suppose you'll drown unless I save you, but I'll drown if I try to save you. That permits me to let you die. Suppose, instead, that I'll die unless I kill you. That doesn't permit me to kill you. The cost to me is the same in both cases: the loss of my life. But that loss justifies me in refusing to save you, but not in killing you. Thus, the presumption against killing you is stronger than the presumption against letting you die.

Let's assume that's true. The Problem of Multiple Threats is fundamentally a problem with discrimination. In this case, the problem of discrimination concerns whom it's permissible to defensively kill. While each account of defensive liability has some plausibility when the number of unjust threats is small, their accounts of liability justifications extend to multiple threats and

---

<sup>133</sup> Jonathan Quong and Joanna Mary Firth, "Necessity, Moral Liability, and Defensive Harm," *Law and Philosophy* 31 (6) (2012): 673-701, at 674.

<sup>134</sup> McMahan, *Killing in War*, 94. Cf. McMahan, "Liability, Proportionality, and the Number of Aggressors," at 19.

<sup>135</sup> James Rachels famously held that the former was true, whereas others have denied it. See James Rachels, "Active and Passive Euthanasia," in Peter Singer (ed.) *Applied Ethics* (New York: Oxford University Press, 1986), 29-35, at 32. For a reply, see Warren S. Quinn, "Actions, Intentions, and Consequences: The Doctrine of Doing and Allowing," *The Philosophical Review* 98 (3) (1989): 287-312.

appear problematically “undefeated,” as Gordon-Solmon says.<sup>136</sup> When a liability justification to defensively harm an unjust threat is undefeated, it’s all-things-considered permissible to harm them. Thus, the accounts of liability targeted by the Problem of Multiple Threats are problematic *because* of whom they say we’re permitted to defensively harm. This is particularly true of the Rights Account and the Responsibility Account, since the Culpability Account endorses liability justifications only for *culpable* threats, and there’s less agreement that killing multiple culpable threats is problematic. Still, insofar as it’s assumed that there is a stronger moral presumption against killing than against letting die, a view delivering the wrong judgment in the former case is less plausible than a view delivering the wrong judgment in the latter case, at least with respect to discrimination.<sup>137</sup>

Here’s the rub: Nearly everyone rejects the Pacifist View because it appears to discriminate wrongly. That’s why nearly everyone accepts the Pacifist View’s alternative, the Standard View. But the Standard View appears to discriminate wrongly, too, as the Problem of Multiple Threats shows.<sup>138</sup> What’s more, since we should think wrongly killing is worse than wrongly letting die, we should (other things being equal) prefer the theory that endorses wrongful letting die over the theory that endorses wrongful killing. That means endorsing Pacifist View over the Standard View. While this doesn’t show that the Standard View is incapable of resolving the Problem of Multiple Threats, it does underscore the seriousness of its advocates doing so.

---

<sup>136</sup> Gordon-Solmon, “Self-Defense Against Multiple Threats,” at 126.

<sup>137</sup> This isn’t meant to be a controversial claim. It’s a typically acceptable form of moral analysis to think that between any two moral views, *a* and *b*, where *a* endorses *x* as permissible and *b* endorses *y* as permissible, that: If *x* is a worse moral wrong than *y*, then it’s less implausible to endorse *y* as permissible than *x*, and thus also less implausible to endorse *b* over *a* as the right moral view.

<sup>138</sup> Someone might object that it’s not liability but *undefeated* liability that’s problematic for the Standard View, at least with respect to the Problem of Multiple Threats. First, the same is arguably true of Pacifism. If pacifists denied the existence of liability justifications for defensive harming but substituted some other justification for it, Pacifism wouldn’t be as counterintuitive as most currently think. Second, it’s the *justificatory* function of liability generating the Problem of Multiple Threats, and thus it *is* liability that’s problematic.

### 3. Three Recent Defenses of the Standard View

In this section, I explore and evaluate a variety of recent proposals for mitigating the Problem of Multiple Threats, offering reasons to think none of the recent proposals succeeds and that the problem is undefeated. This leaves the Standard View immune to a serious discrimination problem with respect to liability to defensive harm and, as I argued in the last section, a *worse* discrimination problem than faced by Pacifism.

#### 3.1. Gordon-Solmon's Defense

Kerah Gordon-Solmon's recent and influential paper, "Self-Defense Against Multiple Threats," attempts to solve the Problem of Multiple Threats by appealing to wide proportionality. Her solution is meant to absolve the Responsibility Account, but as we'll see it also bears helpful fruit for the Rights Account and the Culpability Account.

Gordon-Solmon focuses her attention on the *Multiple Conscientious Drivers* case. If Pedestrian kills all 1,000 drivers, Gordon-Solmon maintains the pedestrian "would also cause a non-trivial side-effect harm to each one," which explains the widespread intuition that "the badness of the harms *to the drivers* is excessive, or disproportionate weighed against the good of saving the pedestrian."<sup>139</sup> However, because the harm to each of the 1,000 doesn't exceed the harm they pose to the pedestrian, relative to the drivers' degree of responsibility, the excessive nature of the

---

<sup>139</sup> Gordon-Solmon, "Self-Defense Against Multiple Threats," at 128-129.

harm they suffer doesn't count against narrow proportionality. Indeed, the non-trivial side-effect harms to which Gordon-Solmon appeals are unavoidable side effects of inflicting the narrowly proportionate harm of killing each driver. These harms can, however, "be aggregated in wide proportionality calculations" since the drivers aren't liable to them.<sup>140</sup>

What are the non-trivial side-effect harms that Pedestrian imposes on each of the 1,000 drivers? Since the threat to Pedestrian's life is overdetermined, given that each driver will kill her, it's permissible for Pedestrian to kill any particular driver only if she kills *all* of the 1,000 drivers, since otherwise killing any of the drivers won't be instrumental in saving her life. Since instrumentality is a necessary condition of liability, there's a liability justification for defensively killing driver 1,000 only if Pedestrian kills drivers 1-999, since Pedestrian will be killed unless all 1,000 drivers are killed.<sup>141</sup> By killing any given driver, therefore, Pedestrian "satisfies a necessary condition for every other driver's liability." Once she's killed drivers 1-999, she "thereby secures the liability justification on the basis of which she kills driver 1,000."<sup>142</sup> Gordon-Solmon then remarks:

Thus, despite the air of paradox, the pedestrian's killing all of the drivers *also* makes each of them substantially worse off as a side-effect, by bringing it about (1) that each is liable to be killed, and (2) that, consequently, she kills each. I submit that being made substantially worse off in this way constitutes a non-trivial harm.<sup>143</sup>

---

<sup>140</sup> Gordon-Solmon, "Self-Defense Against Multiple Threats," at 129.

<sup>141</sup> Gordon-Solmon, "Self-Defense Against Multiple Threats," at 126 and 129.

<sup>142</sup> Gordon-Solmon, "Self-Defense Against Multiple Threats," at 130.

<sup>143</sup> Gordon-Solmon, "Self-Defense Against Multiple Threats," at 130.

The harm of being made liable to defensive harm is, she thinks, a non-trivial harm. Thus, it's a harm that counts toward wide proportionality calculations. Since it counts *non-trivially* toward those calculations, it makes the defensive killing of high numbers of minimally responsible threats widely disproportionate and therefore (other things being equal) morally impermissible.

Gordon-Solmon is cautious here, claiming that her solution works for minimally responsible threats but not culpable threats. She claims that the "impersonal badness" of securing a liability justification "decreases as a threat's culpability increases." As a result,

[A]s the threats' culpability increases, increasingly high numbers need to be harmed in the relevant way for those harms to violate wide proportionality. I suspect the relevant harms to *fully* culpable threats (who are highly blameworthy for threatening great harm) either aren't impersonally bad at all, or aren't bad enough to clear the threshold of additivity.<sup>144</sup>

Jeff McMahan dissents from Gordon-Solmon, claiming that "it seems arbitrary to suppose" that the harm of having one's liability justification secured doesn't count "just because those would suffer them are culpable rather than minimally responsible."<sup>145</sup> Consequently, McMahan holds that Gordon-Solmon's solution "seems to imply that it is also impermissible to kill the culpable killers," which he takes to be implausible.<sup>146</sup> It's worth noting that some, such as Frowe, don't regard McMahan's preferred view (i.e., that it's permissible to kill countless culpable threats) as plausible, whereas others do.<sup>147</sup>

---

<sup>144</sup> Gordon-Solmon, "Self-Defense Against Multiple Threats," at 132.

<sup>145</sup> McMahan, "Liability, Proportionality, and the Number of Aggressors," at 18.

<sup>146</sup> McMahan, "Liability, Proportionality, and the Number of Aggressors," at 18.

<sup>147</sup> Frowe, *Defensive Killing*, at 209. For someone who concurs with McMahan, see Draper, *War and Individual Rights*, at 118-119.

Another important criticism is offered by McMahan. It concerns the *circumstantial* versus the *agential* conditions of liability. On their Responsibility Account, the agential condition of liability is being morally responsible for an unjust threat. Circumstantial conditions include considerations like instrumentality. If, for example, Pedestrian is unable to kill all 1,000 drivers, then the circumstances are such that Pedestrian can't save her life and, thus, none of the drivers is liable to be killed by her. On Gordon-Solmon's view, killing drivers 1-999 doesn't secure their agential liability; they've already done that by being morally responsible for unjust threats to Pedestrian. Rather, killing drivers 1-999 secures their circumstantial liability by removing the impediment to instrumentality, since instrumentality is a necessary condition for liability. On McMahan's view, once all of the drivers satisfy the agential conditions of liability, "[t]he removal of an impediment to the circumstantial conditions of liability is not a relevant cost."<sup>148</sup> He considers a case in which 1,000 minimally responsible threats will kill Victim, but only one threat is in the room. To prevent the others from getting to her and overdetermining her death, Victim harmlessly prevents them from getting to her. In doing so, Victim removes the impediment to the circumstantial conditions to liability, thereby securing the one threat's liability. But, as McMahan observes,

[T]he fact that the removal of the one responsible killer's moral shield against liability would be worse for him does not seem to count *at all* against preventing the others from arriving. It is not a cost that weighs against the threatened harm to the innocent victim in the assessment of wide proportionality.<sup>149</sup>

---

<sup>148</sup> McMahan, "Liability, Proportionality, and the Number of Aggressors," at 16.

<sup>149</sup> McMahan, "Liability, Proportionality, and the Number of Aggressors," at 16-17.

McMahan's analysis is plausible, but it's open to Gordon-Solmon to offer the following reply: By harmlessly preventing the 999 other threats from getting to her, Victim secures the circumstantial liability conditions for *only one* threat. What's troublesome about the Problem of Multiple Threats, according to Gordon-Solmon, is the *additivity* of the side-effect harms: They hold enough weight to count against wide proportionality *en masse*, but not in small numbers. Thus, we should expect the side-effect harm of securing the circumstantial conditions for one threat's liability to count for very little, or perhaps nothing.<sup>150</sup>

Another problem with Gordon-Solmon's analysis is that it seems unable to explain the seeming impermissibility of killing 1,000 unjust threats *simultaneously* and *instantly*.<sup>151</sup> It's easy to see how killing threats 1-999, one at a time, imposes the non-trivial side-effect harm on the threat 1,000, but it's more difficult to see how the harm can be imposed if everyone is instantly killed, for then there's no one to whom the relevant harm can attach. Suppose, for example, that Pedestrian has the power instantly to kill all 1,000 by snapping her fingers and that she does so. Each threat's death is temporally simultaneous with the securing of each threat's liability, and thus each threat is dead at the time of their liability's actualization. Yet it seems difficult to imagine how the side-effect harm of having one's liability secured is a harm to any of the threats, much less the sort of harm that counts toward wide proportionality calculations. Thus, Gordon-Solmon's explanation of the impermissibility of killing multiple threats appears restricted to cases in which

---

<sup>150</sup> There are unrelated cases where it's unclear whether a moral consideration counts for extremely little or for nothing. For example, consider Hereth's example in which one's friend wants her ceramic penguin to face North, but the penguin has been filled with explosives that will kill the friend unless the penguin is destroyed. In this case, do your friend's wishes count for nothing or for very little? It's difficult to say, but neither answer seems utterly implausible. See Hereth, "Against Self-Defense," at 618.

<sup>151</sup> McMahan considers an example like this in "Liability, Proportionality, and the Number of Aggressors," at 26, but offers a different objection: namely, that defensively killing all 1,000 drivers simultaneously self-secures the justification for defensively killing them. Thus, he thinks a kind of bootstrapping is at work in accounts like Gordon-Solmon's.

some threats are killed before others. Such an explanation appears too narrowly applicable and thus misses the heart of the Problem of Multiple Threats.

### 3.2. McMahan's Defense

Jeff McMahan has offered a solution to the Problem of Multiple Threats. He had previously offered other solutions, but he now regards them as implausible. His new solution is to motivate a new kind of proportionality requirement, which (following McMahan) I'll call *aggregate proportionality*.

To motivate this, McMahan asks us to consider again the *Multiple Conscientious Drivers* case. McMahan claims "we can judge intuitively that killing all thousand [drivers] would be disproportionate."<sup>152</sup> This, he thinks, is an intuitive judgment that doesn't depend on prior theoretical commitments about whether instrumentality and narrow proportionality are internal or external to liability. To say that they're internal to liability is to say they're necessary conditions for liability. On this view, each of the drivers is liable to defensive harm only if harming them would be instrumental in saving Pedestrian's life and narrowly proportionate. On the other, external view of liability, instrumentality and narrow proportionality aren't necessary conditions for each driver's liability, but are rather external constraints to all-things-considered permissibility. In other words, each of the driver's can be liable to defensive harm even if harming them would be either non-instrumental or narrowly disproportionate. What harming each of the drivers *wouldn't* be is all-things-considered permissible, because instrumentality and narrow proportionality function not as litmus tests for liability, but rather as litmus tests for all-things-

---

<sup>152</sup> McMahan, "Liability, Proportionality, and the Number of Aggressors," at 25.

considered permissibility. Recall that what it means for someone to be liable to some harm is that they wouldn't be wronged by that harm. But even if they wouldn't be wronged by that harm, it might nevertheless be wrong to inflict that harm on them. For example, someone might be liable to being killed but killing them would also kill dozens of non-liable parties nearby, in which case it would be all-things-considered impermissible to kill the liable party.

Because our intuitions about the disproportionality of killing all 1,000 drivers in *Multiple Conscientious Drivers* doesn't assume the internal or external view of liability, McMahan claims that the kind of proportionality we have in mind isn't narrow proportionality. This is because we judge intuitively that killing all 1,000 drivers is disproportionate *prior to* determining whether each is liable. What's more, McMahan claims, we don't have wide proportionality in mind either, because our intuitive judgment depends partially on the degree of responsibility of each driver—a consideration that makes no difference with respect to wide proportionality. This is why McMahan claims that most people think it's all-things-considered permissible to kill 1,000 *culpably* responsible drivers but not 1,000 *minimally* responsible drivers. Thus, since we intuit killing all 1,000 minimally responsible drivers is disproportionate prior to settling each driver's liability, and since this intuition depends partly on the responsibility of the 1,000 drivers, we can reasonably appeal to a sense of proportionality distinct from narrow and wide proportionality. Because it concerns the numbers, it's about proportionality in the aggregate (i.e., narrow proportionality).<sup>153</sup> Call this first argument for aggregate proportionality the Pre-theoretical Argument. In addition to this argument, McMahan makes another:

---

<sup>153</sup> McMahan, "Liability, Proportionality, and the Number of Aggressors," at 25-26.

If [the Pre-theoretical Argument] is right, it resolves the circularity problem mentioned [earlier]. That was that it seemed that whether it is permissible to kill all thousand responsible killers depends on whether each is liable, which depends on whether killing each would be effective, which in turn depends on whether the others will be killed, which depends, assuming compliance with morality, on whether it is permissible to kill them all. But whether it is permissible to kill all thousand does *not* depend on whether each is liable. We can know that it is *impermissible*, because disproportionate in the aggregate, without knowing whether each is liable.<sup>154</sup>

Call this second argument for aggregate proportionality the Anti-circularity Argument. What this argument is designed to do is to show that attempts to resolve the Problem of Multiple Threats by appeals to the individual liability of each driver results in vicious circularity, which means narrow proportionality (on which each driver's individual liability depends, on the internal view, and on which the all-things-considered permissibility of killing each driver depends, on the external view) is endemic to the circularity. Thus, once we cordon off wide proportionality,<sup>155</sup> only aggregate proportionality can avoid the circularity problem. This doesn't mean that wide and narrow proportionality aren't requirements for permissible defense; it means, instead, that there's an additional requirement.<sup>156</sup>

---

<sup>154</sup> McMahan, "Liability, Proportionality, and the Number of Aggressors," at 25.

<sup>155</sup> McMahan nowhere considers that wide proportionality, unlike narrow proportionality, has the (perhaps minimal) theoretical virtue of solving the circularity problem, since if it's widely disproportionate to kill all 1,000 drivers, then it's all-things-considered permissible to kill them regardless of whether each is liable. Yet this seems like a small oversight, as he thinks (for reasons discussed above) that wide proportionality isn't at issue in the Problem of Multiple Threats.

<sup>156</sup> McMahan doesn't mean to commit himself to some form of consequentialism since he maintains that narrow proportionality is a requirement for permissible defense, which consequentialists would deny. As I argue below, however, McMahan's proposal runs the risk of collapsing into a kind of functional consequentialism.

David Rodin makes two remarks on McMahan's proposal. The first is that McMahan's proposal is theoretically more demanding than Rodin's own proposal. According to Rodin, all 1,000 drivers (whether minimally responsible or culpable) are liable to defensive harm but that there's a lesser evil obligation to refrain from harming them. Thus, on Rodin's view, defensively harming all 1,000 drivers wouldn't wrong the drivers, but it would be wrong.<sup>157</sup> Since McMahan's proposal is theoretically more demanding, there's reason to look elsewhere—for instance, to Rodin's solution.<sup>158</sup>

Rodin's second remark is that McMahan's view doesn't as well accord with "our pre-theoretical intuitions that the liability of individuals to harm ought not to be affected by the number of aggressors."<sup>159</sup> Recall: McMahan's view isn't that the number of aggressors defeats their individual liabilities. He does think liability affects the all-things-considered permissibility of defensively harming a given number of aggressors in the following way: The threshold for permissible defensive killing is *higher* for minimally responsible threats than for culpably responsible threats, and the difference in threshold is explained by their differing degrees of responsibility. But this suggests Rodin gets things backwards: Aggregate proportionality doesn't affect individual liability; rather, individual liability affects aggregate proportionality.<sup>160</sup>

There's a more fundamental problem with McMahan's solution: It entails too restrictive a view of defensive liability, at least by the lights of commonsense the Standard View. For McMahan, as for Gordon-Solmon, how much harm someone is liable to depends both on the harm they pose and how responsible they are for posing it. The general thesis is that someone's liability

---

<sup>157</sup> The details of Rodin's view will be discussed in section 4.3.

<sup>158</sup> David Rodin, "The Lesser Evil Obligation," in Saba Bazargan-Forward and Samuel C. Rickless (eds.) *The Ethics of War* (New York: Oxford University Press, 2017): 28-45, at 44.

<sup>159</sup> Rodin, "The Lesser Evil Obligation," at 44.

<sup>160</sup> In his reply to Rodin, McMahan denies this charge. See McMahan, "Liability, Proportionality, and the Number of Aggressors," at 29.

increases with their responsibility, which is tantamount to saying that someone's moral interest in not being harmed is *discounted* in proportion to their degree of responsibility. This is why, on McMahan's view, the interests of fully culpable threats are heavily discounted, perhaps even *fully* discounted, given their strong degree of culpability. The heavy discounting of their interests explains why it's all-things-considered permissible to kill swaths of them: because each of their moral interests in not being harmed is discounted to an extent that, even collectively, doesn't amount to much, if anything. But if that's true, then the interests of minimally responsible threats should be *barely* discounted; that is, their interest should be discounted at a rate scarcely higher than non-responsible threats, whose interests aren't discounted *at all* on McMahan's Responsibility Account. That means killing minimally responsible threats, unlike killing fully culpable threats, will violate aggregate proportionality very quickly—even when there are so few as *two* of them. Consider a variant of *Multiple Conscientious Drivers* in which only two conscientious drivers threaten Pedestrian. On the Responsibility Account, Pedestrian isn't liable to defensive harm at all, and nor is there anything else that discounts Pedestrian's moral interest against harm. But the drivers are liable since they meet the threshold of responsibility, albeit minimally. Thus, each driver's interests are discounted only slightly more than Pedestrian's. If that's true, then the aggregate interests of two such drivers count for nearly twice as much as Pedestrian's interests.<sup>161</sup> While this changes nothing about each driver's individual liability, it does affect aggregate proportionality: the proportionate weighing of respective interests. It appears all it takes to make things disproportionate in the aggregate sense is to have two minimally responsible

---

<sup>161</sup> The sense in which their interests “count” has the same functional sense as it does when McMahan and others claim that the interests of fully culpable threats are “discounted.” The interests of the latter, even in the aggregate, don't violate aggregate proportionality because they're heavily (if not fully) discounted, and thus aren't disproportioned to the contrastingly weightier interests of Pedestrian. But that isn't true of minimally responsible threats, whose interests (despite being somewhat discounted, given their liability) count for nearly as much as Pedestrian's.

threats. Yet that's very counterintuitive, at least with respect to commonsense the Standard View.<sup>162</sup> This is partially why McMahan and Gordon-Solmon, when they attempt to show the Problem of Multiple Threats, appeal to examples like *Multiple Conscientious Drivers* in which a very high number of minimally responsible threats must be killed to save Pedestrian, because what seems counterintuitive about killing 1,000 minimally responsible threats is the sheer number of them. Killing ten or five or two minimally responsible threats, by contrast, doesn't seem like too many, yet it would be under McMahan's aggregate proportionality proposal. Thus, aggregate proportionality lowers the threshold of permissible self-defense too far by the commonsense standards of the Standard View, and so should be rejected as a solution to the Problem of Multiple Threats.

### 3.3. Rodin's Defense

The third and final solution to be discussed is David Rodin's. He begins by discussing lesser evil justifications for non-consensual harming. These justifications are distinct from liability justifications for non-consensual harming. With liability justifications, "we can explain how it can be permissible to [harm] a person *because* of facts about their rights," whereas with a lesser evil justification, "we can explain how it can be permissible to treat a person in a certain way *despite* facts about their rights."<sup>163</sup> Consider, for example, a famous case in tort law in which a farmer "justifiably destroys his neighbor's field to stop a wildfire from engulfing a town."<sup>164</sup> The neighbor still had a right against others not to burn his field, but his right was "'overridden' or 'outweighed'

---

<sup>162</sup> This isn't a problem for Pacifism. The intuitive instability of McMahan's position is that it entails the permissibility of defensively killing one, *and only one*, minimally responsible threat. Pacifism doesn't claim this.

<sup>163</sup> Rodin, "The Lesser Evil Obligation," at 29.

<sup>164</sup> Rodin, "The Lesser Evil Obligation," at 29.

by considerations that are independent of the underlying grounds for the possession of [his] right.”<sup>165</sup>

Rodin then explains lesser evil *obligations*. These are obligations that, like their counterpart lesser evil justifications, explain deontic moral facts (like obligations and permissions) not because of facts about the harmed person’s rights, but because of independent considerations. He offers an example of this:

Imagine that you are subject to an unjustified assault that would likely leave you with a painful, but impermanent, injury. Suppose you could avert this injury by breaking the assailant’s arm, and that this treatment would be narrowly proportionate in the circumstances (it would not be excessive in relation to the assailant’s liability to be harmed). However, the assailant is a manual laborer whose large family is entirely dependent on him. If he breaks his arm he will lose his income and the family will be destitute, moreover there is no evidence that breaking his arm will prevent the assailant or any other person from committing assault in the future. It is plausible that you would have an all things considered obligation not to engage in self-defense in these circumstances—despite the fact that you clearly have the right of self-defense with respect to the assailant (you would not wrong him or infringe *his* rights by breaking his arm).<sup>166</sup>

It’s worth carefully unpacking what happens in this detailed case. First, your assailant is liable to defensive harm, and therefore lacks a right that you not (defensively) harm him. That’s why you wouldn’t be liable to the assailant if you broke his arm. Second, you would nevertheless act wrongly if you

---

<sup>165</sup> Rodin, “The Lesser Evil Obligation,” at 31.

<sup>166</sup> Rodin, “The Lesser Evil Obligation,” at 32.

broke the assailant's arm, given how your action adversely affects third parties. As a result, says Rodin, you "ought to waive" your right to defensive action.<sup>167</sup> Since it's true that you *ought* to waive your rights, there's an all-things-considered moral obligation to refrain from defensive action.<sup>168</sup> Third, this is all true despite the fact that you have a right to defensively break the assailant's arm. As Rodin phrases it,

In the current cases the right bearer is all things considered obligated to behave in a certain way toward a person despite the fact that they have the right with respect to that party to behave otherwise.<sup>169</sup>

Lesser evil considerations can generate a duty to sacrifice a good that one has a right to possess, by requiring an agent to abstain from self-defensive acts he has a liberty to perform.<sup>170</sup>

Thus, on Rodin's view, the all-things-considered obligation not to break the assailant's arm doesn't imply that you lack a right to do so. Rather, it implies you have an obligation to waive your defensive right—an impossibility if you didn't still possess that right.

---

<sup>167</sup> Rodin, "The Lesser Evil Obligation," at 32.

<sup>168</sup> The claim that Defender "ought to waive" his right is somewhat confusing. Rodin isn't entirely clear about which right is being waived: the (claim-)right against unjust harm or the (liberty-)right to self-defend against that unjust harm. If it's the former, then Defender isn't *wronged* when he's killed, since he waives his right against being killed, and surely that isn't what Rodin means. If, on the other hand, he waives only his (liberty-)right to self-defend against unjust harm in a particular circumstance, it's unclear what moral work is done by the waiving. Suppose, for example, that Defender refuses to waive his right to self-defense. Would that make it permissible for him to self-defend? It seems not, on Rodin's view, and it certainly doesn't make it permissible for Defender to self-defend if he *does* waive the right to self-defense. But then Defender's obligation not to self-defend is maintained *whether or not* he waives the right to self-defend, and thus it's not clear why he "ought to waive" that right. Perhaps Rodin merely means that Defender ought not to exercise that right, in which case the right to self-defend persists but goes unused.

<sup>169</sup> Rodin, "The Lesser Evil Obligation," at 32-33.

<sup>170</sup> Rodin, "The Lesser Evil Obligation," at 36.

From here, Rodin moves to solve the Problem of Multiple Threats. He claims, first, that the deaths of “impersonal harms are aggregated for the purposes of proportionality in lesser evil reasoning.”<sup>171</sup> Rodin thinks the impersonal badness of someone’s death isn’t mitigated or eliminated by their liability, such that it can be impersonally bad to harm someone who’s liable to that harm.<sup>172</sup> This is true, Rodin thinks, even for culpable threats. According to him, while some might think culpable threats are “both liable to, and deserving of, harm in a self-defense situation,” but they are mistaken because “the notion of deserving harm is suspect” and “has no role to play in a theory of self-defense.”<sup>173</sup> Thus, on Rodin’s view, even the deaths of culpable threats count as impersonally bad and can be calculated into lesser evil deliberations. At some point, Rodin claims, they count enough:

[At] some point the aggregated impersonal harm inflicted on the aggressors must exceed the prospective good by such a margin that the right to kill them becomes subsumed.<sup>174</sup>

The prospective good in question is the right of the victim. When that good will be secured only if a very high number of liable threats are defensively killed, securing the good is outweighed by the impersonal badness of the threats’ deaths. Said another way, killing all of the threats—for instance, all 1,000 drivers—secured a good by causing an outweighing evil. When that happens, the right of the victim to self-defend is maintained but trumped by other considerations, making it all-things-considered impermissible to kill all of the threats.

---

<sup>171</sup> Rodin, “The Lesser Evil Obligation,” at 37.

<sup>172</sup> Cf. Rodin, “The Lesser Evil Obligation,” at 36. This doesn’t mean that the death of *any* liable threat is impersonally bad. Some liable threats might not be harmed by their deaths. All Rodin claims is that someone’s *liability* to be harmed doesn’t make their being harmed any less impersonally bad.

<sup>173</sup> Rodin, “The Lesser Evil Obligation,” at 37. Rodin’s view therefore conflicts with mine, which is that culpable threats deserve defensive harm. I briefly argue for this position in section §2.

<sup>174</sup> Rodin, “The Lesser Evil Obligation,” at 37.

Rodin is clear that a lesser evil obligation is generated only when the “aggregated impersonal value of the aggressors’ lives...exceed the impersonal value of the putative defender’s life by a very significant margin.”<sup>175</sup> Like McMahan, Rodin thinks “the impersonal badness of their deaths is not discounted as steeply compared with fully liable culpable aggressors.”<sup>176</sup> For Rodin, this means that the lesser evil obligation is generated at a lower number for minimally responsible threats than for culpable threats. For example, if the lesser evil obligation kicks in at 1,000 for minimally responsible threats, it won’t kick for culpable threats until a higher number is reached—1,500, for example. The lesser evil obligation, then, permits you to kill very few minimally responsible threats, but far more fully culpable threats. Helen Frowe appears to endorse this solution. She writes,

Roughly, this explanation will have to hold that there comes a point at which the aggregate value of the lives of the attackers, even responsible attackers—at perhaps at some point even culpable attackers—is such that one cannot bring about their deaths.<sup>177</sup>

Frowe doesn’t explicitly endorse the lesser evil strategy, and nor does she offer a fully developed account. Nevertheless, the bare bones of her solution bear a striking resemblance to Rodin’s more developed account.

One problem with Rodin’s solution is that it won’t work for culpable threats. The reason why is because culpable threats are not only liable to being killed defensively; they also, contrary

---

<sup>175</sup> Rodin, “The Lesser Evil Obligation,” at 38.

<sup>176</sup> Rodin, “The Lesser Evil Obligation,” at 39.

<sup>177</sup> Frowe, *Defensive Killing*, at 209.

to what Rodin claims, *deserve* to be killed defensively, since desert supervenes on culpability.<sup>178</sup> In other words, if you culpably threaten me with death, then it's true not only that killing me is unfair but also that killing you is fair. It's fair precisely because, in turning the lethal threat against you, you'd be getting your just deserts. (This alone won't make it all-things-considered permissible to kill you since, as Jeffrey Reiman observes, the fact that Threat deserves to be killed is consistent with it being all-things-considered impermissible to kill him.<sup>179</sup>) It's impersonally *good* for individuals to get their just deserts, and so it's impersonally good for culpable threats to be defensively killed.<sup>180</sup> This matters because, unlike certain other accounts of who's liable to defensive harm, the Culpability Account can't appeal to the impersonal badness of the deaths of culpable threats to explain the wrongness of killing multiple culpable threats. Thus, the Culpability Account is in a somewhat worse position to explain away the Problem of Multiple Threats than the Responsibility Account. Like the latter, it can't explain the wrongness in terms of liability. But unlike the latter, it can't explain the wrongness in terms of the impersonal badness of the threats' deaths. Rodin insists there's a lesser evil obligation not to kill vast numbers of culpable threats.

---

<sup>178</sup> As McMahan points out, liability is distinct from desert insofar as liability is "essentially *instrumental*" (e.g., in saving Victim's life) whereas desert isn't essentially so. (See McMahan, "Liability, Proportionality, and the Number of Aggressors," at 8.) Thus, saying culpable threats deserve to be defensively killed might appear to imply that it's all-things-considered permissible to kill them even after the threat to Victim has abated. If so, the Culpability Account is in extremely rough waters. However, it's possible to claim that one's culpable responsibility for unjustly killing someone *diminishes* (or, at least, importantly changes) after that individual is dead, since one is no longer threatening the individual. This would mean, for example, that the unjust killer deserves different treatment in the context of *preventing* an unjust killing than in the context of (for example) *punishing* the unjust killing, such that there's room to say (1) that multiple culpable threats deserve to be killed and (2) that they don't deserve to be killed once they're no longer *threats* (e.g., once the unjustly threatened individual is dead). Defenders of the Culpability Account could also maintain that although desert isn't essentially instrumental (as McMahan claims), it can in some contexts have instrumental constraints, such as constraints after the culpable killing has been done.

<sup>179</sup> Jeffrey Reiman, "Why the Death Penalty Should Be Abolished in America," in *The Death Penalty: For and Against* (Lanham, MD: Rowman & Littlefield, 1998): 67-132, at 67.

<sup>180</sup> McMahan (*Killing in War*, at 8) holds this view. Some retributivist accounts, like certain Kantian accounts, explain the goodness of getting one's just deserts in an intrinsically personal way: i.e., as a way of respecting agency. On such an account, killing them is personally good (since it respects their agency) but could also be impersonally bad (since it harms the agents). But on such an account, the personal goodness of respect for agency *trumps* the impersonal badness of harming the agents. Said another way, doing right *by* agents trumps doing bad *to* agents. Since the trumping relation holds for each individual agent, it also holds in the aggregate.

But there's no room for evil if they're culpable: Killing them doesn't wrong them (because they've forfeited their rights) and their deaths aren't impersonally bad (because they deserve them). Rodin denies that culpable threats deserve to be killed or otherwise harmed, but his denial is flat-footed as he fails to motivate this claim.

Another problem with Rodin's lesser evil obligation proposal is that it's susceptible to a threshold problem similar to the one afflicting McMahan's aggregate proportionality solution. Let's begin with an important question: What evils are suffered by the unjust threats? It isn't the transgression of their rights, because Rodin claims the relevant threats are liable and so have forfeited those rights. Presumably, then, it's the loss of their impersonally valuable lives. But what's the numerical threshold at which the lesser evil obligation applies? There are only two possible numerical thresholds: (a) the threshold at which the disvalue of many deaths "overrides" or "outweighs" a right against (unjust) death, and (b) the threshold at which the disvalue of many deaths "override" or "outweigh" the disvalue of one death. I take it that the threshold for (a) is higher than (b).

How high is the threshold for (a)? Whatever it is, if it's high enough to "override" or "outweigh" Defender's right against unjust harm, then it's permissible for third parties to *kill* Defender on the same basis. (After all, Defender's right is *trumped*.) Yet that seems to imply that there's a lesser evil justification—perhaps even a lesser evil *obligation*—for killing Defender. Suppose that each of the 1,000 minimally responsible drivers know this,<sup>181</sup> and suppose they also know that either Defender will kill all of them or they will kill Defender,<sup>182</sup> and suppose that they

---

<sup>181</sup> I have chosen minimally responsible threats because, unlike some culpable threats, they lack a prior intention to kill Defender.

<sup>182</sup> It might matter *how* Defender will kill them. In the examples under consideration, Defender kills them by intentionally *using* a weapon (e.g., by blowing up their cars with an anti-tank gun). But we can imagine cases in which Defender kills them by unintentionally *being* a weapon (e.g., by representing a threat to them in the same way a crossing deer represents a threat to a car, where the sheer force of the bodily impact will kill the driver).

purpose to kill Defender to secure the lesser evil. If they do this, then *they have a lesser evil justification (indeed, a lesser evil obligation) for killing Defender*, and thus aren't liable to defensive harm. The reason why is because having a justification for posing a harm entails that one isn't posing an unjust harm; or, at the very least, it entails that one isn't *liable* to be harmed on the basis of posing a justified harm.<sup>183</sup> This is why, on Rodin's view, it's impermissible to defensively harm the farmer who justifiably prevents the destruction of a town by burning his neighbor's field.<sup>184</sup> While this has the upshot of explaining why it's impermissible to kill multiple threats, it explains only why it's impermissible to kill multiple *justified* threats and not multiple *unjust* threats, and it's the latter group at the heart of the Problem of Multiple Threats.

The second possibility is (b): the threshold at which the disvalue of many deaths "override" or "outweigh" the disvalue of one death. How high is the threshold for (b)? For minimally responsible threats, it's roughly two-to-one: two minimally responsible threats are just shy of having *twice* the interests of Defender.<sup>185</sup> If that's what Rodin means, then the lesser evil obligation kicks in at *two* threats, which seems absurd. What's more, it's the same over time: You can save your life from unjust death only twice, and no more. For these reasons, we should reject Rodin's solution to the Problem of Multiple Threats.

#### **4. An Intractable Problem**

---

<sup>183</sup> Jeff McMahan defends this view at length in Jeff McMahan, "Self-Defense Against Justified Threateners," in Helen Frowe and Gerald Lang (eds.) *How We Fight* (New York: Oxford University Press), 104-136, at 118-123. For a different view, see Adam Hosein, "Are Justified Aggressors a Threat to the Rights Theory of Self-Defense?" in *How We Fight*, 87-103. However, the dispute between McMahan and Hosein is relevant only if the multiple threats have a lesser evil justification but not (as I've claimed) a lesser evil obligation to kill Defender.

<sup>184</sup> Rodin, "The Lesser Evil Obligation," at 29.

<sup>185</sup> On Rodin's view, the threshold for culpable threats is more than two-to-one. See Rodin, "The Lesser Evil Obligation," at 38-40.

I have offered reasons to believe none of the existing proposals solves the Problem of Multiple Threats. But that's merely a *contingent* reason for accepting Pacifism, as another proposal might come along that succeeds. In this section, I shall argue that what generates the Problem of Multiple Threats is commitments essential to the Standard View. This offers strong reason to believe that the Standard View is incapable of avoiding the problem, which (other things being equal) offers a reason to believe Pacifism.

#### **4.1. Avoiding *Ad Hoc* Proposals**

One necessary criterion of success for any proposal is that it not set an *arbitrary limit* on the number of unjust threateners one may permissibly kill in self-defense. This means that *pairwise comparisons* should hold across numerically analogous cases.<sup>186</sup> For example, consider

##### *Five Threats*

Defender is lethally attacked by five unjust threats. Defender can save his life only if Defender kills all five of them.

Any plausible variant of the Standard View will say it's permissible for Defender to kill all five unjust threats. Now, consider

##### *Ten Threats*

---

<sup>186</sup> See McMahan, "Liability, Proportionality, and the Number of Aggressors," at 5.

Defender and Friend are lethally attacked by ten unjust threats. Because Friend is unconscious, only Defender can protect them against the unjust threats. Defender can save their lives only if Defender kills all ten of the threats.

If it's permissible for Defender to defensively kill five unjust threats in defense of zir own life, then it should also be permissible for Defender to defensively kill ten unjust threats in defense of zir own life *and* someone else's life. After all, it would be permissible for Friend to kill five of the unjust threats, and there's nothing importantly morally different about Defender doing so on Friend's behalf.<sup>187</sup> Said another way: If saving one life justifies killing five unjust threat, then saving two lives should justify killing twice that many (i.e., ten).

#### 4.2. Ratios

The avoidance of *ad hoc* proposals appears to require a commitment to *ratios*. This partially justifies McMahan's proposal, which appeals to proportionality. It might also justify Rodin's proposal which, despite not appealing to proportionality, appeals to lesser evils. It goes without much saying that, for Rodin, the evil must be *enough* of a lesser evil to generate an obligation not to defensively kill liable parties.

It would be uncharitable to the Standard View to assume the ratios must be precise. As a result, I won't assume avoiding *ad hoc* proposals entails that level of precision. But there's a

---

<sup>187</sup> Consider the following variant of *Five Threats*:

*Five Threats\**. Friend is lethally attacked by five unjust threats but can't defend themselves. Fortunately, Defender is nearby and can save Friend's life only by killing all five threats.

Another commitment of the Standard View is that it's permissible for Defender to engage in other-defense on behalf of Friend, killing all five unjust threats. Notice, however, that *Ten Threats* is the numerical conjunction of *Five Threats* and *Five Threats\**.

difference between not requiring precise ratios and not having just any requirements for the ratios proposed. Not just anything goes.

Returning to the prior cases of *Five Threats* and *Ten Threats*, we saw that the number of unjust threats it's permissible to kill in the latter case is *approximately double* the number of unjust threats it's permissible to kill in the former. Said another way: If it's permissible for Defender to defensively kill ( $n$ ) number of unjust threats to prevent them from killing zir, then it's permissible for Defender to defensively kill *approximately* ( $n \times 2$ ) unjust threats to prevent them from killing Defender and Victim. As I've explained it, this appears to be an unavoidable commitment of doubling the number of potential victims. This generalizes, and the following principle expresses the more generalized truth:

#### RATIO PRINCIPLE

It's permissible for Defender to defensively kill unjust threats in approximate proportion to the number of people who would otherwise be unjustly killed by the unjust threats.

So understood, the principle is agnostic about *what* the ratio is. It implies only that the number of people Defender is permitted to defensively kill must be approximately proportioned to the number of victims. This is a combination of two distinct claims, each of which is plausible. The first claim is that as the number of would-be murder victims grows, so does the number of unjust threats it's permissible to kill. Second is the claim that these are approximately proportioned. Both of these are things we should expect, assuming the Standard View. It's extremely plausible, on the Standard View, to think that even if Defender is forbidden to kill a million unjust threats to save zir life, Defender is permitted to do so to save the lives of a million people (including, perhaps, zirself).

It's also plausible to suppose, as the examples of *Five Threats* and *Ten Threats* show, that there's some kind of rough proportionality consideration at work in both cases.

### 4.3. The Problem Resurfaces

Acceptance of the RATIO PRINCIPLE preserves and formalizes the basic commitments of the Standard View. Unfortunately, it merely returns us to the Problem of Multiple Threats, albeit a somewhat mitigated one. I'll first explain why the Problem of Multiple Threats is a two-pronged problem.

The most obvious specter raised by the Problem of Multiple Threats is *limitless liability justifications*. As McMahan phrases the problem,

Even more troublingly, if there is a liability justification in each pairwise comparison that is unaffected by the number of killers, there is then *no* limit to the number of killers it would be permissible to kill in defense of a single victim.<sup>188</sup>

As we've seen, that's indeed concerning. While it's perhaps the deepest problem revealed by the Problem of Multiple Threats, it isn't the only problem. Consider the following example:

#### *Million Limit*

Defender is besieged by a million lethal, unjust threats and has a liability justification for killing each of them. However, there's a non-arbitrary cut-off for liability justifications at one million.

---

<sup>188</sup> McMahan, "Liability, Proportionality, and the Number of Aggressors," at 6.

In the example, Defender is permitted to kill a million lethal, unjust threats, but not one more. For the sake of argument, we're supposing this limit is non-arbitrary. Even if it were non-arbitrary, it seems that Defender is still permitted to kill *far* too many unjust threats in *Million Limit*. Thus, the Problem of Multiple Threats is not simply problematic when there are unlimited liability justifications. It's also problematic when there are a *finitely oversized* number of liability justifications. I'll now defend the view that, on the Standard View, the number of liability justifications will *necessarily* be (finitely) oversized, unless the ratio endorsed is approximately 1:1. As argued in the previous section, it's doubtful the ratio can be that low if the Standard View is plausible. The reason why is because, if the ratio were that low, it would imply Defender isn't permitted to defend against more than one unjust threat, which seems implausible.

How high *is* the ratio? It's useful here to see the numbers given in the literature. Most provide high numbers to solicit counterintuitive results. For example, McMahan's example involves ten prisoners and five-hundred guards: a ratio of 1:50.<sup>189</sup> Gordon-Solmon's example involves one pedestrian and one-thousand drivers: a ratio of 1:1,000.<sup>190</sup> Each of these authors believes these are excessive, and thus also implicitly accept that the ratio must be smaller than this. I shall assume the very modest thesis that *the ratio is no smaller than 1:5*, since it seems implausible in cases like *Five Threats* that Defender isn't permitted to kill all five unjust threats. Having said this, consider the following example:

### *Five Thousand*

---

<sup>189</sup> McMahan, *Killing in War*, at 24.

<sup>190</sup> Gordon-Solmon, "Self-Defense Against Multiple Threats," at 128.

Defender and 999 other non-labile parties will be killed by 5,000 unjust threats unless Defender kills all of them. Defender can accomplish this.

Two things are important to note here. First, the numbers conform to the ratio of 1:5. Second, the number of liability justifications seems excessive. Killing five thousand people to save one thousand seems excessive. Or, if you don't think that's excessive, consider increasingly larger variants of the 1:5 ratio, such as

*Five Million*

Defender and 999,999,999 other non-labile parties will be killed by 5,000,000,000 unless Defender kills all of them. Defender can accomplish this.

Again, the numbers conform to the ratio. Killing *four million more people* than you save seems wildly excessive. I shall spare the reader a discussion of further iterations of this ratio, as the point should now be clear: Mere conformity with a ratio is insufficient to solve the Problem of Multiple Threats, as is conformity with a very plausible ratio. And matters become far worse the larger the ratio becomes.

There's a further, more decisive argument for the conclusion that a ratio as low as 1:5 is problematic, at least for some variants of the Standard View. Consider the case of

*Wartime*

To save the lives of 1,000,000 innocent (i.e., non-labile) people, Soldier must kill 5,000,000 innocent people. Soldier does so.

It's undisputed that Soldier's actions are impermissible because they clearly violate wide proportionality. Killing the 5,000,000 innocents, in Gordon-Solmon's words, is "excessive, weighed against the good" it would achieve.<sup>191</sup> The good it achieves (i.e., saving 1,000,000 innocents) is indeed a great good, but still a *lesser* good to leaving 5,000,000 innocents alive. Finally, in *Wartime*, I explicitly say that none of the parties are liable to be killed. Thus, whatever the criteria for liability, none of the 5,000,000 has it. Their interests in not being killed are robust and 'full'; they are not discounted whatsoever. Furthermore, it's clearly impermissible to kill the 5,000,000. Now consider a variant on *Wartime*:

*Traffic*

To save the lives of 1,000,000 trapped on Interstate 5, Patrolwoman must kill 5,000,000 minimally responsible drivers who will otherwise run them over.

Because the drivers are minimally responsible for the unjust threat they pose to the 1,000,000 trapped pedestrians, their interests are discounted somewhat. On the Responsibility Account, the 5,000,000 minimally responsible drivers are liable to be killed by Patrolwoman. That's a problem. The interests of minimally responsible threats are discounted only *slightly* more than the interests of non-liable parties—just enough to make them liable, but not so discounted that they're the liability equivalents of fully culpable threats. Let's assume that, in *Wartime*, Soldier is permitted to kill no more than 999,999,999 non-liable individuals to save 1,000,000.<sup>192</sup> Yet, on the Responsibility Account, Patrolwoman is permitted to kill 4,000,000 more individuals than she

---

<sup>191</sup> Gordon-Solmon, "Self-Defense Against Multiple Threats," at 125.

<sup>192</sup> This is a very liberal estimate. For reasons I won't explore here, it's far from clear whether Soldier is permitted to kill so many individuals to save a net of one individual.

saves. This *quadruples* the interest-discounting of minimally responsible threats relative to non-liable individuals, which far exceeds the slight discounting that should occur between non-liable individuals and minimally responsible threats.<sup>193</sup> Thus, killing 5,000,000 minimally responsible drivers in *Traffic* isn't permissible. Because that judgment violates the RATIO PRINCIPLE, the Responsibility Account is false.<sup>194</sup>

All of this is very bad news for the Standard View. To summarize: The Standard View permits Defender to kill at least five unjust threats. If that's permissible, then so is killing in conformity with the implicit ratio. But *some* killing, even if it's in conformity with the implicit ratio, is impermissible. This is a problem for the Standard View, since it appears unable to avoid the Problem of Multiple Threats. While this doesn't perhaps show that the Standard View is *false*, it does offer strong reason to reject it.

## References

Alexander, Larry, and Kimberly Ferzan. (2009) *Crime and Culpability: A Theory of Criminal Law* (New York: Cambridge University Press).

---

<sup>193</sup> To be clear: This is true despite the fact that killing the 4,000 in *Wartime* violates wide proportionality, since they're non-liable, whereas killing the 4,000 in *Traffic* violates narrow proportionality, since they're liable to be harmed but not that much.

<sup>194</sup> In itself, this is a significant conclusion. The Responsibility Account is the predominant view of moral liability. However, versions of this argument apply to the Culpability Account and the Rights Account.

Anscombe, Elizabeth. (1981) "Mr. Truman's Decree," in *The Collected Philosophical Papers of G.E.M. Anscombe: Volume III: Ethics, Religion, and Politics* (Oxford: Wiley-Blackwell): 62-71.

Doggett, Tyler. (2011) "Recent Work on the Ethics of Self-Defense," *Philosophy Compass* 6 (4): 220-233.

Draper, Kaila. (2016) *War and Individual Rights* (New York: Oxford University Press).

Draper, Kaila. (2009) "Defense," *Philosophical Studies* 145 (1): 69-88.

Ferzan, Kimberly. (2012) "Culpable Aggression: The Basis for Moral Liability to Defensive Killing," *Ohio State Journal of Criminal Law* 9: 669-697.

Frowe, Helen. (2014) *Defensive Killing* (New York: Oxford University Press).

Gordon-Solmon, Kerah. (2017) "Self-Defense Against Multiple Threats," *Journal of Moral Philosophy* 14: 125-133.

Hereth, Blake. (2017) "Against Self-Defense," *Social Theory and Practice* 43 (3): 613-635.

Hosein, Adam. (2014) "Are Justified Aggressors a Threat to the Rights Theory of Self-Defense?" in Helen Frowe and Gerald Lang (eds.) *How We Fight* (New York: Oxford University Press), 87-103.

Kaufman, Whitley. (2010) "Self-Defense, Innocent Aggressors, and the Duty of Martyrdom," *Pacific Philosophical Quarterly* 91: 78-96.

McMahan, Jeff. (2017) "Liability, Proportionality, and the Number of Aggressors," in Saba Bazargan and Samuel C. Rickless (ed.) *The Ethics of War* (New York: Oxford University Press): 3-27.

McMahan, Jeff. (2016) "The Limits of Self-Defense," in Christian Coons and Michael Weber (eds.) *The Ethics of Self-Defense* (New York: Oxford University Press), 185-210.

McMahan, Jeff. (2014) "Self-Defense Against Justified Threateners," in Helen Frowe and Gerald Lang (eds.) *How We Fight* (New York: Oxford University Press), 104-137.

McMahan, Jeff. (2009) *Killing in War* (New York: Oxford University Press).

McMahan, Jeff. (2005) "The Basis of Moral Liability to Defensive Killing," *Philosophical Issues* 15 (1): 386-405.

McMahan, Jeff. (1994) "Self-Defense and the Problem of the Innocent Attacker," *Ethics* 104: 252-290.

Narveson, Jan. (2003) "Terrorism and Pacifism: Why We Should Condemn Both," *International Journal of Applied Philosophy* 17 (2): 157-172.

Narveson, Jan. (1965) "Pacifism: A Philosophical Analysis," *Ethics* 75 (4): 259-271.

Otsuka, Michael. "Killing the Innocent in Self-Defense," *Philosophy and Public Affairs* 23 (1) (1994): 74-94.

Rachels, James. (1986) "Active and Passive Euthanasia," in Peter Singer (ed.) *Applied Ethics* (New York: Oxford University Press): 29-35.

Reiman, Jeffrey. (1998) "Why the Death Penalty Should Be Abolished in America," in *The Death Penalty: For and Against* (Lanham, MD: Rowman & Littlefield), 67-132.

Rodin, David. (2017) "The Lesser Evil Obligation," in Saba Bazargan-Forward and Samuel C. Rickless (ed.) *The Ethics of War* (New York: Oxford University Press): 28-45.

Ryan, Cheyney. (1983) "Self-Defense, Pacifism, and the Possibility of Killing," *Ethics* 93 (3): 508-524.

Thomson, Judith Jarvis. (1991) "Self-Defense," *Philosophy and Public Affairs* 20 (4): 283-310.

Thomson, Judith Jarvis. (1971) "A Defense of Abortion," *Philosophy and Public Affairs* 1 (1): 47-66.

Quinn, Warren S. (1989) "Actions, Intentions, and Consequences: The Doctrine of Doing and Allowing," *The Philosophical Review* 98 (3): 287-312.

Quong, Jonathan, and Joanna Mary Firth. (2012) "Necessity, Moral Liability, and Defensive Harm," *Law and Philosophy* 31 (6): 673-701.

## Against Self-Defense

**ABSTRACT:** The ethics of self-defense is dominated by the Standard View, which claims that at least some cases of self-defensive assault are permissible. I shall defend the radical view that the Standard View is false. My argument proceeds as follows. Every permissible act of self-defensive assault has the modal property being possibly massively overdetermined. Such ‘super-permissible’ acts of assault are ones in which agents are objectively permitted to perform those acts in morally trivializing or cavalier fashion: i.e., agents need not ‘think twice’ about inflicting or permitting harm; they are permitted to assault persons as if it were morally insignificant. Yet this is never true, since assaulting persons is always morally serious. It follows that there are no acts of permissible self-defensive assault.

### 1. Introduction

This essay concerns what I shall call the *Standard View* of self-defensive assault:

Standard View:

Some instances of self-defensive assault are morally permissible.

By ‘assault’, I mean the sort of physical or psychological harm characteristically inflicted in self-defense.<sup>195</sup> Here is one useful account:

You assault me at time  $t$  only if and because:

- (i) You intentionally and directly cause an event, C;
- (ii) prior to C, a fully-informed me would have had conclusive prudential reason to escape C (absent some other C-type circumstance if I succeeded in escaping C) if I aimed to protect myself; and

---

<sup>195</sup> For another account, one that characterizes these acts not as acts of assault but as acts of *violence*, see Audi (2009). Audi notes that violence can be done accidentally, but accidents are not the sorts of things there can be a moral obligation to avoid, and the sort of pacifism I defend is principally contrasted with the Standard View in virtue of disagreements about the sort of harmful act which is always and everywhere wrong.

(iii) I do not escape C.

On this account, several essential features of assault are present.<sup>196</sup> The first condition identifies two: *intentionality* and *directness*. Any unintentional harm is not assault; there can be no accidental assault. Assault is therefore purposeful, intentional. Regarding directness: if you hire an assassin to kill me, then *you* do not assault me. At most, you intend my assault. Thus, assault requires causal directness. There are degrees and kinds of directness, however, but the directness condition is designed primarily to exclude as instances of assault cases in which some agent (*qua* agent) distinct from you harms me.

I am *fully-informed*. Consider a case in which, prior to C, I have conclusive prudential reason to escape C, but this conclusive prudential reason is grounded in some misleading epistemic consideration: You aim a gun at me but don't intend to fire. This would hardly constitute assault. Thus, I must (again, prior to C) have conclusive prudential reason to escape C, where this conclusive prudential consideration holds on the assumption that I am fully-informed.

The circumstance, C, must itself be *counterfactually sufficient to motivate conclusive prudential reason to avoid C*. Consider a case in which I attempt to escape C (say, by running out the back door of the pub) but, when attempting to escape, finds myself in another C-type circumstance (say, there is a violent gang outside ready to mutilate me). Staying in the bar and being shot in the leg might, from my standpoint, be prudentially *better* than attempting to escape being shot in the leg.<sup>197</sup> Still, this would not make shooting me in the leg a case of non-assault.

---

<sup>196</sup> I exclude force as a requirement, since it is not obvious to me that force is a necessary condition of assault.

<sup>197</sup> Perhaps this is why consensual euthanasia does not seem to be a case of assault. If it is in the person's best interest to die and if the person consents, it seems to many that killing the person (or, on passive variations, intentionally letting the person die) is compassionate and hardly an instance of providing prudential reason to avoid C. Where it is best for someone not to escape but rather to die, it is not an act of assault to kill her (or, on passive variations, intentionally to let her die).

Circumstance C is one for which I have conclusive prudential reason to escape *absent* some alternate C-type circumstance.

Escaping C must be for the sake of *me*. Someone might have conclusive prudential reason to move their car from the train tracks because, otherwise, someone else will intentionally and directly smash it to bits. But this would not be assaulting me because it is not an act against me, even if it might exert causal effects on me.

Finally, that which I have conclusive prudential reason to avoid must *happen*. After all, I might have conclusive prudential reason to avoid C (where C is shooting me in the leg); there might be no other pressing C-type circumstance (no gang outside); and still I might escape C. Upon doing so, I will have avoided assault while meeting the other conditions. Thus, the prior conditions are not sufficient for assault. It is necessary that I fail to escape C for the account to be completed.

Widespread acceptance is enjoyed by the thesis that self-defensive assault of this kind is sometimes morally permissible. Thomson (1991: 283) says that many “think of self-defense as morally transparent. What could be clearer than that morality permits a person to save his or her life against threats to it?” She then asks her readers to consider the case of Villainous Aggressor:

[Y]ou are standing in a meadow, innocently minding your own business, and a truck suddenly heads toward you. You try to sidestep the truck, but it turns as you turn. Now you can see the driver: he is a man you know has long hated you. What to do? You cannot outrun the truck. Fortunately, this is not a pure nightmare: you just happen to have an antitank gun with you, and can blow up the truck. Of course, if you do this you will kill the driver, but that does not matter: it is morally permissible for you to blow up the truck, driver and all, in defense of your life. (Thomson 1991: 283)

Thomson is explicit that this moral permission is not mere excuse, but that blowing up the truck (and the driver) is morally permissible, justified killing.<sup>198</sup> This is because

...blowing up the truck in Villainous Aggressor is not something you ought not do. We cannot plausibly say that you ought not blow up the truck, but will only be in a measure at fault, or in no measure at fault, for doing so: you simply *may* blow up the truck. Morality permits it. (Thomson 1991: 283)

Cases like Villainous Aggressor are widely regarded as paradigmatic for justified instances of self-defensive assault. Jonathan Quong explicitly endorses Thomson's conclusion about Villainous Aggressor and other cases Thomson raises, writing, "Like Judith Jarvis Thomson, I think these are all cases where it is permissible to kill one person in order to save my own life" (2009: 507). Quong is hardly alone in this endorsement. Many philosophers endorse Thomson's conclusion about Villainous Aggressor and structurally similar cases, and in their comments suggest consensus:

In these circumstances, few of us would condemn you for killing in self-defense. Nor would we condemn a third party who intervened on your behalf by killing your attacker. (Draper 1993: 73)

---

<sup>198</sup> Thomson's claim that "it does not matter" if the driver is killed is perhaps best understood as: it does not matter, all things considered (which is compatible with, but does not entail, the stronger claim that the death of the driver is altogether morally irrelevant).

Draper's talk of condemnation appears to support the view that defenders are in many cases unworthy of condemnation because they act rightly.

...the simple point is that if somebody else is intent on leaving you dead, and there is strong immediate evidence that the only really feasible way to prevent this is to leave him dead first, then normal people prefer the latter to the former. (Narveson 2003: 158)<sup>199</sup>

It is widely accepted, for example, that if one person, Albert, culpably threatens the life of another person, Betty, then he may have forfeited his rights against serious harm being imposed on him in Betty's defence. (Quong and Firth 2012b: 674)

Still others, such as Jeff McMahan (1994c: 193)<sup>200</sup>, Michael Otsuka (1994: 74), and Helen Frowe (2014: 1) claim that this view is widely held. If these judgments about *Villainous Aggressor* cases are true, then self-defensive assault is at least sometimes justified.<sup>201</sup> Beyond agreement about *Villainous Aggressor* cases, there is even broader agreement that self-defense is permissible. Draper (2009: 69) refers to self-defense as "perhaps the most widely recognized justification for inflicting harm." Tyler Doggett (2011: 220) makes the broad claim that "[t]he literature on self-defense agrees that killing" *Villainous Aggressor* is permissible. Quong (2012: 45) endorses the Standard View unequivocally: "Sometimes it is morally permissible to seriously harm or kill people in self-defense, or in defense of others."

---

<sup>199</sup> Narveson's use of "normal people" is suggestive of consensus.

<sup>200</sup> McMahan goes further than some in suggesting that self-defensive assault is just as often obligatory as it is merely permissible. Still, McMahan is a clear instance of a philosopher who endorses the permissibility of self-defensive assault.

<sup>201</sup> Draper and Frowe make similar judgments in Draper (1993, 73) and Frowe (2010, 247).

Some philosophers even contend that the permissibility of self-defensive assault in some cases is either itself obvious or is itself entailed by obviously true moral principles. Among them are Seth Lazar (2009: 700), McMahan, F.M. Kamm (2012: 219), and David Rodin:

Defensive rights seem to be entailed in a very basic way by rights *to* things. Thus if I have a right to *X*, then it seems to follow as a simple corollary that I have the right to take measures to prevent my right to *X* from being violated. (2002, 37)

McMahan (1994b: 252) goes so far as to say that the Standard View is an *axiom* of contemporary ethical theory. Concurring with McMahan that ordinary morality includes the possibility of justified self-defensive assault, Lazar (2009: 700) claims that self-defense can remove “*the whole wrong* in killing a person” (emphasis mine). This suggestion makes theories of self-defensive attractive tools for those who believe that war can be justifiably waged. Rodin’s powerfully simple observation that a right to something entails a right to preserve it supports the permissibility of self-defensive assault: if one’s right to life is (wrongly) threatened, one is permitted to exercise one’s right to do what is necessary to protect that right, including in cases where killing or seriously harming others is necessary to protect one’s right to life.

I shall argue that the Standard View is false. My argument will proceed as follows. If it’s ever permissible to harm an individual with the sort of moral status most human beings are believed to have, that permissibility can be overdetermined. Furthermore, it can be massively overdetermined. Thus, every permissible act of self-defensive assault has the modal property *being possibly massively overdetermined*. Where the permissibility of harming a being is massively overdetermined, I say it is “super-permissible.” I then argue super-permissible acts of assault are

ones in which agents are objectively permitted to perform those acts in morally trivializing or cavalier fashion. Among other things, this means that agents need not ‘think twice’ about inflicting or permitting harm; they are permitted to assault persons as if it were, morally speaking, insignificant.<sup>202</sup> Yet this is never true, since assaulting others is always deeply morally serious. It follows that there are no acts of permissible self-defensive assault. Though the focus of this paper will address *self-defensive* assault, it should be reasonably clear that a similar argument could be given against other kinds of assault (e.g., other-defensive or punitive).

## 2. Permissible Action and the Cavalier Constraint

I claim that there is a *Cavalier Constraint* for permissible action. This constraint is a side-constraint in the sense described by Robert Nozick: an act is permissible for some agent only if it does not violate any (absolute) side-constraint.<sup>203</sup> This constraint is as follows:

### Cavalier Constraint:

---

<sup>202</sup> My argument is therefore distinct from Andrew Fiala’s (2014) argument against the Standard View. Fiala’s argument is that the Standard View gets its moral support by avoiding hard cases and focusing instead on sterilized ‘Trolley’ cases in which it appears obvious that self-defensive assault is justified. Because such cases abstract away from more difficult cases, Fiala claims the Standard View gets its moral support too cheaply. Fiala’s argument is an important one, but it’s not the one I want to make. The sense in which killing—or, to use my favored term, ‘assault’—becomes problematically *cavalier* is the sense in which its performer would be objectively justified in acting as if assaulting someone was no big deal, given the host of considerations that make the act permissible. My worry is that the Standard View implies that this is possible, which I use as a *reductio ad absurdum* against the Standard View. There are, then, two possible approaches here: The first is Fiala’s, which points out that killing people often gets messier than we expect, and we can never assume that killing is as sterilized as ‘Trolley’ cases make it appear. The second is that regardless of how things *actually* go when we kill people, the Standard View implies that we can kill people as if it’s no big deal, and that’s false. The difference between Fiala’s argument and mine is that Fiala thinks defenders of the Standard View get more than they bargain for even if they’re right about the sterilized cases, whereas I think they’re wrong about even the sterilized cases. My thanks to an anonymous reviewer for noting the similarities between my argument and Fiala’s and for pressing my to clarify how my argument is distinct from his.

<sup>203</sup> See Nozick (1974: 28-33).

There is some class of actions that can't be permissibly treated cavalierly, where an action  $\phi$  is treated cavalierly if there is some possible world in which some agent is objectively morally permitted to treat or regard the doing of  $\phi$  as if it were morally "no big deal."

What I shall show in this section is that the following premise is true:

Premise 1: The Standard View is true only if there is some instance of permissible self-defense that could not, in principle, violate the Cavalier Constraint.

Imagine that a moral agent comes permissibly to regard the performance of some action as if it were morally inconsequential. The sense in which this agent acts permissibly is objective, and not (merely) epistemic<sup>204</sup>: Her action is made permissible by some moral consideration, and she knows this. Suppose that she regards  $\phi$ -ing in a morally cavalier way by treating it with a sort of moral ease: She very casually does  $\phi$  and doesn't give a second thought to doing so.<sup>205,206</sup>

The Cavalier Constraint doesn't apply to *all* actions. It remains open that there are actions that *can* be done cavalierly. For example, suppose  $\phi$ -ing is the act of respecting the wishes of my friend to leave her favorite ceramic penguin facing North, but imagine that someone has filled the penguin with explosives that will kill my friend unless I knock it off the shelf, destroying it and

---

<sup>204</sup> Cf. Jeff McMahan's (2011: chapter 4, esp. 162-163) overview of the distinction.

<sup>205</sup> "Regarding" might appear tacitly to endorse some form of doxastic voluntarism, since it requires agents neither to treat nor regard a violation or disregarding of  $\phi$  as morally inconsequential. But the sense of regarding need not be doxastic, and the only relevant doxastic cases are ones in which one *does* have even some voluntary control over one's beliefs. For example, if one voluntarily selects actions one knows will result in believing, or believing more strongly, that  $\phi$  is morally inconsequential, then one would act wrongly in choosing those actions.

<sup>206</sup> Cf. Nagel (2005). Nagel suggests that citizens have an interest in having political leaders who are ruthless and carry out the business of the state, but their ruthlessness should not become internalized. (In the terms employed here, ruthless actions should not be carried out with ease.)

averting the blast. In such a case, failing to  $\phi$  is not only *permissible*; it's so very permissible that I would be objectively justified in knocking the ceramic penguin off the shelf without a second thought.

Whereas some actions can be permissibly done in cavalier fashion, however, others can't. To use just one example of a morally grave action, consider torture. Let  $\phi$  be the act of not torturing someone and assume, for the sake of argument, that acting *against*  $\phi$  can be permissible, perhaps on grounds of self-defense or other-defense, as Uwe Steinhoff argues:

There is no morally relevant difference between self-defensive killing of a culpable aggressor and torturing someone who is culpable for a deadly threat that can be averted only by torturing him. (Steinhoff 2006: 337).<sup>207</sup>

In some theories of self-defense, it's permissible to assault unjust aggressors when necessary for self-protection because those aggressors have, via their unjust aggression, *forfeited* their right not to be assaulted. Though many advocates of self-defense take issue with Steinhoff's claim that self-defensive assault is morally indistinct from self-defensive torture, assume that self-defensive torture is sometimes permissible.<sup>208</sup> Consider the following case:

*Uneasy Kai* knows that self-defensive torture is permissible and finds herself in a circumstance in which torturing a person is their only means of self-defense. They torture the threat but without moral ease: They do not enjoy it, they're cautious to inflict only the

---

<sup>207</sup> Cf. Steinhoff (2012: 19-26) and Steinhoff (2013: chapter 9).

<sup>208</sup> See the challenge to this conclusion by Bufacchi and Arrigo (2006).

harms they deems necessary, and participates with a high degree of hesitation and trepidation.

Uneasy Kai does not enjoy torture in any way; they regard it as unfortunate that it is at all instrumentally necessary for their own defense (or the defense of others). Kai realizes too that torture, even in self-defense, is a very morally serious activity, and they responds appropriately by taking great pains to be sure that each cut and each blow is absolutely necessary, giving their threat ample time between each clipped nail to reconsider sharing information relevant to the threat against Uneasy Kai.<sup>209</sup> But consider a variation on Uneasy Kai:

*Nonchalant Kai* learns that Bibi is nearby, and they decides to kidnap Bibi and torture him for information necessary to save Kai's life. With each electrical shock and pried fingernail, nonchalant Kai proceeds with moral nonchalance: They don't give much additional thought to whether some harm they inflicts is necessary or whether their torturous acts are right.

Unlike Uneasy Kai, who responded appropriately to the moral graveness of torture, Nonchalant Kai fails to respond appropriately. Torture is intrinsically grave and the moral risks associated with it are plentiful: inappropriate emotional attitudes towards or reactions to harming others; objectifying victims; resorting to inappropriate means (torturing non-threatening family, for example); torturing victims even after they have surrendered information due to uncertainty about

---

<sup>209</sup> This does not assume that there are no metaphysically vague cases in which it is indeterminate whether inflicting some harm *H* is necessary (or, is among a set of comparable alternatives, one of which is necessary) to achieve goal *G*, or that there are no cases in which it is epistemically unclear whether *H* is necessary to achieve *G*. Rather, the claim is simply that no agent is possibly objectively justified in *casually assuming* that *H* is necessary to achieve *G* when *H* is morally serious.

the truthfulness of their claim to have surrendered the correct information; and more. Thus, even if torturing in self-defense is permissible, this does not permit anyone to treat torture in any way that would constitute a nonchalant or casual attitude toward its moral seriousness. Everyone everywhere always has a duty to take torture seriously, even if they are inflicting it with justification.<sup>210</sup>

Examples like these help to raise two important questions. The first is what makes actions, or action-kinds, subject to the Cavalier Constraint? Second, if those actions or action-kinds cannot permissibly be treated cavalierly, what moral relations and functions (if any) are affected in virtue of the Cavalier Constraint?

It's not necessary for the purposes of this paper to sketch the full set of necessary and sufficient conditions for what qualifies an action for the Cavalier Constraint. It's enough to show that there is a Cavalier Constraint and that it applies to assaulting persons. To this end, one sufficient condition will be identified and defended.

The short defense is as follows. All persons have properties, some of which are essential. One essential property of persons is *being extremely morally considerable*. This property necessarily confers a right to all persons, a right they cannot forfeit and a right that cannot be

---

<sup>210</sup> Cheyney Ryan's pacifism is of the skeptical variety. He maintains that "the proponent of killing cannot produce a single argument for why killing another person is permissible" (1983: 509). He proceeds to argue as follows: There's no plausible sense to be made either of the claim that Aggressor *forfeits* his right to life or of the claim that Victim has a *right* to kill Aggressor, yet one of those is necessary if a defense of the Standard View is to succeed. Thus, the Standard View doesn't succeed. Where my argument bears some similarity to Ryan's is where he discusses the *positive* reasons for being a pacifist, namely, "The pacifist's problem is that he cannot create, or does not wish to create, the necessary distance between himself and another to make the act of killing possible" (1983: 521). Ryan offers the example of shooting a fascist with his trousers undone: It doesn't seem at all wrong to kill *fascists*, but it *does* seem wrong to kill blokes with their trousers down. What the pacifist refuses to do, therefore, is to come to see unjust aggressors as—to borrow a phrase from James Kellenberger—"conscious, knowing obstacles...whose removal or neutralization is a means to one's end" (1987: 136). However, unlike Ryan, my point is *not* that we should refuse to distance ourselves from others to the extent that we can, psychologically, see them as things to be killed. Rather, my point is that we should reject as false any view that would objectively permit us to distance ourselves from others to the extent that we can kill them *as if it were no big deal*. I argue that the Standard View does so, and that the Standard View is therefore false.

countervailed: the right not to be treated carelessly or, when harming them, harming them casually.<sup>211</sup>

The longer defense begins with a defense of the claim that all persons possess that essential property. To begin, consider that every person has a presumptive right against being harmed or killed. This right is absolutely presumptive: it holds absent any forfeiting or countervailing conditions.<sup>212</sup> Said another way, the right against being harmed or killed holds whenever the person being harmed or killed neither forfeits that right nor has that right countervailed. Consider such a case. Bailey is not attacking anyone and there is no moral good worth preserving or evil worth preventing. Thus, there is no moral reason to kill Bailey. If you harm or kill Bailey, you have wronged her, and this is true in all possible worlds in which you harm or kill Bailey under these conditions. By modal implication, Bailey *necessarily* has the right not to be harmed or killed absent forfeiture or countervailing conditions.

As it concerns harming or killing persons, the threshold at which countervailing conditions obtain is high. If a gallon of milk will be spilt unless Bailey is harmed or killed, we ought not to harm or kill her. If Bailey will die unless she tosses Shana under the Greyhound bus, Bailey still ought not to toss Shana under the Greyhound bus. Restrictions on permissible harming and killing are plentiful, even in cases of permissible self-defense.

Moreover, even in cases where countervailing (but not forfeiting) conditions do obtain, it is sometimes the case that there exists some *residual duty* to the person or persons harmed or killed.

McMahan, for example, writes:

---

<sup>211</sup> This is true only if there are rights. It is assumed that there are, but the following judgments about what is permissible and what is not do not depend upon the existence of rights.

<sup>212</sup> Forfeiture conditions operate in the following way: if some circumstance obtains, then person Steve forfeits Steve's right. Countervailing conditions operate in the following way: if some circumstance obtains and is stronger than the relevant right, then Steve's right is permissibly overridden (or trumped).

When one thus permissibly acts against a right, I will say that one *infringes* that right, whereas when one impermissibly does what another has a right that one not do, one *violates* that right. Even though an agent acts permissibly in infringing a right, the victim is nonetheless wronged and may thus be owed compensation.<sup>213</sup>

A paradigmatic example of residual duty is a modified trolley case in which either Steve will lose his leg or five people will die. Assuming it is permissible to switch the track such that Steve loses his leg, it appears true both that Steve had a right not to have his leg taken *and* Steve is owed compensation. When non-combatants are unintentionally harmed as a direct result of tactical bombing or other means of offensive warfare, they too are plausibly owed compensation if such compensation can be provided.<sup>214</sup> More strongly, every person whose right is permissibly infringed in McMahan's sense is owed compensation, other things being equal.

The moral significance of entities like persons, each of whom possesses a necessary presumptive right not to be harmed and a presumptive right to compensation in cases of right-infringement is nearly inestimable. These rights appear to confer an extremely considerable moral status: the rights are possessed presumptively, have a remarkably high countervailing threshold, and are possessed by every person in all possible worlds in which those persons exist. Furthermore, the extremely morally considerable nature of persons is grounded in the *modal* properties possessed, which persons continue to possess even if, for example, their right not to be harmed or killed is forfeited or countervailed in the actual world. Supposing that Bibi's right to life or not to

---

<sup>213</sup> McMahan (2011), 10.

<sup>214</sup> See Brummer (1996).

be harmed is forfeited or countervailed in the actual world, it hardly follows that those rights of Bibi's or their presumptive force are forfeited or countervailed *across worlds*. Even if Bibi forfeits his right not to be assaulted today because he aggresses unjustly against Kai, Bibi does not thereby *absolutely* forfeit his right across possible worlds. Rather, Bibi retains his presumptive right not to be assaulted such that when there is no reason to harm Bibi, then harming Bibi is wrong. These essential properties of persons, therefore, ground the following right of persons: the right not to be treated carelessly or, when harming them, harming them casually.

### **3. Moral Relations Affected by the Cavalier Constraint**

Because of the Cavalier Constraint, certain moral relations are affected. Relations affected by persons with the relevant rights will include, among other things, the *overdetermined* relation. To say that the permissibility of self-defense is overdetermined is to say that there is more than one independent moral consideration whose normative implications entail permissible self-defense. For example, if one's permission to protect one's children implies that one is permitted to self-defend,<sup>215</sup> and if one likewise has an independent permission to protect oneself, and if defending oneself is likewise necessary to prevent catastrophe,<sup>216</sup> then that set of permissions supplies independent moral grounds for permissible self-defense. The permissibility of self-defense is overdetermined because, in each case,  $\phi$ -ing is permissible and it requires self-defense. That is, protecting one's children requires first that I defend myself, as does preventing catastrophe. Thus, the permissibility of  $\phi$ -ing in each case *grounds* the permissibility of self-defense.

---

<sup>215</sup> See Davion (1990: 92); Parfit (2011: 205); and Narveson (1965: 262).

<sup>216</sup> See Steinhoff (2006).

There is conclusive reason to believe that all theories of justified self-defense *in principle* permit overdetermination.<sup>217</sup> That is, the following premise is true:

Premise 2: Any possible instance of self-defense could, in principle, have its permissibility massively overdetermined.

According to all Standard theories of self-defense, self-defense is not absolutely wrong (that is, it is actually permissible in some possible circumstance) and can be motivated (that is, there would be moral enough reason to permit self-defense in the circumstance). Each of these conditions is strictly necessary for overdetermination.<sup>218</sup> If self-defense were *absolutely prohibited*, it would necessarily fail to make any act of self-defense permissible or contribute to its permissibility. Thus, the permission to protect one's children or prevent catastrophe could make self-defense permissible or contribute to its permissibility only if self-defense is not absolutely prohibited.

Moreover, it must be the case that self-defense is possibly motivated. This assumption is not built into the conception of self-defense being actually permissible in some circumstance, since self-defense might be permissible even if there is no moral reason to act in self-defense. The criminal who breaks into someone's home in order to steal a leftover blood sample for a harmless science project might use defensive measures to prevent the vandalism, but it is not clear that there is any *moral* reason to prevent this. The person who finds this activity strange but harmless is hardly lacking in moral reasons-responsiveness when she fails to prevent the robbery.

---

<sup>217</sup> Excluding, of course, eliminativist theories of self-defense like pacifism, which imply that there is no permission, and therefore no right, to self-defensive assault.

<sup>218</sup> Both are likewise necessary for moral *determination* (again, in the grounding sense).

While these observations alone do not conclusively support the conclusion that all theories of self-defense permit overdetermination, they do some work in clearing the way of conceptual obstacles. In particular, they show that two possibilities that necessarily would make overdetermination impossible do not hold for theories of self-defense. This provides some support for the conclusion that overdetermination is (necessarily) possible for theories of self-defense. Fortunately, stronger support can be given.

An extended defense of the possibility of overdetermination includes cases in which overdetermination appears obvious. There are, necessarily, possible cases in which there is more than one moral consideration that normatively implies a permission to self-defend. Returning to our earlier examples, we have the following set of permissions<sup>219</sup>:

- (a) All guardians have a *pro tanto* permission to protect children under their care;
- (b) All innocent persons have a *pro tanto* permission to protect themselves; *and*
- (c) All persons have a *pro tanto* permission to prevent catastrophe at serious cost to themselves (but not to others).

Let's also suppose that the following is true at a given time for some agent:

- (d) Unless Kai self-defends right now, Kai will fail to protect children under their care;
- (e) Unless Kai self-defends right now, Kai will fail to protect themselves; *and*
- (f) Unless Kai self-defends right now, Kai will fail to prevent catastrophe.

---

<sup>219</sup> The following set of permissions can easily be substituted for a distinct set where the permissions for self-defense are more obviously such that there is no lexical priority.

Suppose also that there are no permissions distinct from these ones which would further imply that Kai was permitted to self-defend right now. Thus, the moral considerations which permit Kai to self-defend are *exhausted* by considerations (a)-(f). Furthermore, assume that there is no lexical or temporal priority amongst the permissions. That is, assume that each permission *has the same moral strength* (in terms of being a moral consideration) and *arises at the same time*. When Kai finds themselves in the circumstance in question, there is no permission that enjoys a ‘first-to-ground’ privilege in virtue of some lexical or temporal priority. What this means is that no particular permission *uniquely* grounds Kai’s permission to self-defend right now. Yet, because Kai’s permission to self-defend right now *is* grounded by (a)-(f), it follows that the permission is *overdetermined*.

There is another argument for this conclusion: a best-explanation argument. The permission to self-defend appears stronger in cases in which more is at stake than in cases where less is at stake. If Kai’s failure to self-defend would result in the death of their children and the destruction of the greater Baltimore area, then Kai’s permission to self-defend is plausibly stronger than in a case in which Kai’s failure to self-defend would result merely in their own death. In such a case, Kai’s permission to self-defend appears stronger because of the normative influence of certain counterfactual considerations: namely, that the destruction of Baltimore would result and that would be worse. Thus, the fact that the permission to self-defend appears sensitive to these concerns is further evidence for the possibility of overdetermination.

The explanation acquires further strength when it is recognized that the relevant moral considerations (preventing the death of one’s family and preventing the destruction of Baltimore) are each individually sufficient to motivate at least a *pro tanto* permission to self-defend. Because

of this, each consideration is ordinarily permission-*creating*, and it is difficult to see why it would fail to be permission-*affecting* in these cases.

I shall now argue that overdetermination ruled out by the Cavalier Constraint? That is, the following premise is true:

Premise 3: Any act that could, in principle, have its permissibility massively overdetermined is also an act whose performer would, under those conditions, be permitted to perform that act in a cavalier fashion.

To see why, recall that assaulting persons *cannot* be permissibly done cavalierly.<sup>220</sup> If Kai is to permissibly assault a person, they cannot be acting on the basis of moral considerations that make assaulting the person super-permissible, or so permissible that Kai is permitted to assault the person cavalierly. But if assault can be made permissible and in a way that is overdetermined, a problem arises.<sup>221</sup> Suppose that if Kai is to avoid the horrific leveling of Baltimore, they must torture Bibi. As in a previous case, Kai's partner, children, parents, and extended family live in Baltimore. Kai is in Baltimore interrogating Bibi. Grant lastly that torturing Bibi is permissible and overwhelmingly so. If this is true, then Kai's permission to assault Bibi (which includes *torturing* Bibi) is made super permissible by the relevant set of moral considerations, such that the act assaulting Bibi could be permissibly done cavalierly.

---

<sup>220</sup> It is a necessary truth because, as defended above, it is true in every possible world that if persons exist, they have a presumptive right not to be killed or harmed, which entails a corresponding duty on others not to harm or kill persons absent forfeiture and countervailing considerations. But if some claim is true in every possible world, then it is a necessary truth.

<sup>221</sup> The metaphor of the trump card is illustrated powerfully by Dworkin (1984: 153).

We might tack on a fistful of other permissions that coincide to overdetermine the permissibility of torturing Bibi. In addition to (a)-(f), we might add the following:

(g) All persons have a *pro tanto* permission to keep their promises;

(h) All persons have a *pro tanto* permission to demonstrate meaningful opposition to terrorist activities.

We shall also assume, as before, that the following claims are also true:

(i) Unless Kai self-defends right now, Kai will fail to keep their promises; *and*

(j) Unless Kai self-defends right now, Kai will fail to demonstrate meaningful opposition to terrorist activities.

There might in principle be an infinite number of such permissions, or at least such a swath of them that the permission to assault Bibi is super permissible. Recall the case in which a friend's wish to have her ceramic penguin face North can permissibly be infringed since failing to move the penguin will otherwise result in loss of life or limb. The savior's *pro tanto* duty to act within the bounds of her friend's wishes is vastly overridden by the permissions grounded in saving lives. It is because of this significantly overdetermined permissibility that the savior would act permissibly in refusing to give even a second thought to tossing the penguin. Something similar could occur in some cases of assault, however, in which self-defensive assault can be and is super-permissible. Yet this conclusion is vastly at odds with the existence of a Cavalier Constraint for assaulting persons. Because there is such a constraint and because the possibility of permissible

self-defensive assault entails the possibility of overdetermination, and because such overdetermination would result in trivializing an act for which there is a Cavalier Constraint, we should conclude that *self-defensive assault is not even possibly permissible*.

The argument likewise holds even if one adopts a forfeiture view of permissible self-defense. On a forfeiture view, Bibi *loses* a right, the former presence of which would have entailed a legitimate claim against Kai that Kai not attack Bibi.<sup>222</sup> Now Bibi's right is gone, and with it comes the permissibility of killing him. But the permissibility of killing Bibi can still be overdetermined. For example, Frowe (2014: 3) claims that Bibi forfeits his rights through "involuntary behavior," but there might be several simultaneous instances of voluntary behavior by which Bibi forfeits his rights. Perhaps Bibi unjustly attacks Kai and also unjustly attacks a bystander, each of which would be sufficient for Bibi to forfeit his right, rendering permissible an attack by Kai. In that sort of case, Kai is *further permitted* to attack Bibi; it is more important to attack Bibi than it would be if, for example, Bibi attacked only Bibi. Such reasons can be added until the permissibility of assaulting Bibi becomes super permissible, such that Kai would be objectively justified in assaulting Bibi without a second thought. Because this is impossible, given the Cavalier Constraint, and because forfeiture accounts entail that it is possible, forfeiture accounts do not escape the objection.

To complete the argument against the Standard View, only one premise is required, and it's true by definition:

Premise 4: Performing an act in a cavalier fashion would violate the Cavalier Constraint.

---

<sup>222</sup> See McMahan (2011: 10). For a defense of the forfeiture account, see Lang (2014) and Frowe (2014: 3-4 and 72-73). Notably, Lang (2014: 38) holds that whereas the standard picture of permissible self-defense is one in which the wrongful aggressor *loses* her right not to be killed, forfeiture accounts are accounts of *how* that right is lost.

It follows from this and the three prior premises that the Standard View is false.

#### **4. Relations Affected by the Near-Cavalier Constraint**

The lynchpin of the argument against the Standard View concerns overdetermination. In that argument, it was claimed that any number of distinct moral reasons might normatively affect the permission to self-defend against an aggressor. Because of this, the act of assaulting a person would be trivialized in whichever possible worlds the horde of moral reasons overdetermines the permissibility of assaulting a person. In virtue of this false modal implication, the Standard View is false. More formally:

##### *The Argument from the Cavalier Constraint*

(Premise 1) The Standard View is true only if there is some instance of permissible self-defense that could not, in principle, violate the Cavalier Constraint.

(Premise 2) Any possible instance of self-defense could, in principle, have its permissibility massively overdetermined.

(Premise 3) Any act that could, in principle, have its permissibility massively overdetermined is also an act whose performer would, under those conditions, be permitted to perform that act in a cavalier fashion.

(Premise 4) Performing an act of self-defense in a cavalier fashion would violate the Cavalier Constraint.

---

(Conclusion) The Standard View is false.

What might be said against this suggestion is that the argument itself provides principled reason to reject the very possibility of *cavalier* overdetermination while maintaining the possibility of *non-cavalier* overdetermination. This might appear *ad hoc*, but it need not be. The argument offered here might as well be an argument against the possibility of excessive overdetermination, and one might suppose that this is logically compatible with the truth of the Standard View.<sup>223</sup> Moreover, this suggestion is compatible with agnosticism about what the precise cutoff conditions for normative influence are: which moral considerations are disqualified, to what extent normative influence can occur, and the like. Thus, the objector might concede that the argument above would be compelling if there were not countervailing considerations, such as the objection argument provides. More specifically, the Argument from the Cavalier Constraint would succeed in showing that cavalier overdetermination is made possible by the Standard View *only if* there is no principled way for the Standard View to rule out cavalier overdetermination. But that is precisely what is made possible by the argument against the Standard View.

Beyond trivializing the act of assaulting persons, there is a secondary constraint for permissible action. Consider the following case:

*Almost Nonchalant Kai* captures and tortures Bibi to prevent a catastrophe. Kai realizes that torturing Bibi cannot be permissibly done cavalierly, and they do not demonstrate a total lack of moral attentiveness to whether their actions are necessary, right, etc. They are

---

<sup>223</sup> Unless, as has been argued here, and as will be further argued in a moment, the Standard View does *entail* the possibility of cavalier overdetermination.

not quite cavalier. But they are *close* to being cavalier: They remain very comfortable, assuming, and regard torturing Bibi as *nearly* inconsequential.

While not acting quite as wrongly as Nonchalant Kai, Almost Nonchalant Kai *does* act wrongly. They act wrongly because torturing a person is not near-trivial, and thus it is wrong to treat torturing a person as if it is near-trivial. Thus, for some acts, there is a *secondary* constraint:

Near-Cavalier Constraint:

There is some class of actions that can't be permissibly treated near-cavalierly, where an action  $\phi$  is treated *nearly* cavalierly if there is some possible world in which some agent is objectively morally permitted to treat or regard the doing of  $\phi$  as if it were *nearly* morally "no big deal."

To forbid agents from treating or regarding some act as nearly morally inconsequential is, again, not to require those agents to feel or believe a particular way about that act. Furthermore, it is logically possible for some agent to violate her duty to observe the Near-Cavalier Constraint for the acts admitting of such a constraint, yet fail to do so in a blameworthy way. An agent might be unaware that there is such a constraint, and thus would not plausibly be blameworthy for failing to abide by the normative demands of the constraint. Thus, the agent might be epistemically (or subjectively) morally justified in treating or regarding  $\phi$  as if it were morally inconsequential, but she could not be *objectively* morally justified.

As in the case of torture, I claim that there is a Near-Cavalier Constraint for the act of assaulting persons,  $\phi$ ; thus,  $\phi$  cannot be permissibly done in a *nearly* cavalier fashion. It would be

objectively wrong for an agent to assault a person by breaking her finger in a morally near-cavalier way, say, by taking it seriously but in a borderline way: one is very near *not* taking it seriously. Because a Near-Cavalier Constraint exists for  $\phi$ , which includes all cases of assaulting persons, it extends to cases in which assaulting persons is done under widely-believed permitting conditions for self-defense.

The argument from the Cavalier Constraint proceeded as follows: There is a Cavalier Constraint for  $\phi$ , which entails that  $\phi$  cannot be permissibly done cavalier, and  $\phi$  could be permissibly performed cavalierly *if* severe overdetermination for permissibly doing  $\phi$  was possible. Necessarily, if the Standard View is true, then severe overdetermination for performing  $\phi$  is possible. Therefore, the Standard View is false.

The Cavalier Constraint for assaulting persons draws a line in the sand: There is a cap on the number of permissions or extent to which those permissions can normatively affect the permissibility of an act (e.g., assaulting a person). For example, suppose as before that the permissions to protect your children, yourself, and prevent catastrophe are each sufficient to permit you to assault a person in self-defense. If those considerations permitted this, they would severely overdetermine and therefore *render possibly cavalier* assaulting that person. However, suppose that if only the first two considerations permit assault, then the assault would *not* be trivialized by severe overdetermination. In that case, the third consideration (i.e., preventing catastrophe) *would be* a permission absent the other considerations, but it doesn't function as a permission here because it can't do so.<sup>224</sup>

---

<sup>224</sup> Thus, acting so as to prevent catastrophe is a *pro tanto* permission in the following sense: it is normative or has normative implications, other things being equal. Absent special conditions, agents are permitted to prevent catastrophe. Alternately or in addition, acting so as to prevent catastrophe is a *pro tanto* permission in the following sense: agents are permitted to prevent catastrophe, but not all avenues to prevent catastrophe are permissible (e.g., genocide).

<u>Consideration</u>	<u>Permission-Affecting Status</u>
<i>Prevent catastrophe</i>	non-affecting
CAVALIER CONSTRAINT	
<i>Protect your children</i>	affecting
<i>Protect yourself</i>	affecting

This is analogous to a case in which the permission to prevent catastrophe does not imply that one can permissibly commit genocide: Because genocide is absolutely wrong, it cannot be morally permitted, and thus no reason can stand in the relevant permitting relation. The Cavalier Constraint makes severe overdetermination impossible, and therefore makes impossible any and all relations that would entail such severe overdetermination.

The objector's move is to make possible *some* overdetermination with respect to assault, but not enough to render it permissible to do cavalierly. If successful, this move would for example effectively allow certain moral considerations to permit assault, but would not permit another consideration from simultaneously doing so. It is helpful to think of the constraints as thresholds, and Figure 1 (above) is helpful in illustrating the relations.

Above the constraint, which is rightly conceived and effectively illustrated as a threshold, no permissions normatively affect  $\phi$ . Beneath the constraint threshold, moral considerations can normatively affect  $\phi$ , even overdetermining the permission for doing  $\phi$ .<sup>225</sup> Because of the Near-Cavalier Constraint, however, our scale must include a secondary threshold, as represented in Figure 2.

---

<sup>225</sup> Because specifying where the threshold 'begins' is a matter of specifying the total set of sufficient conditions for rendering an act permissible to perform cavalierly, and because this essay focuses on only one sufficient condition, no more will be said about the finer details of the constraint in regards to a scale as represented in Figure 1 (above).

<u>Consideration</u>	<u>Permission-Affecting Status</u>
<i>Prevent catastrophe</i>	non-affecting
CAVALIER CONSTRAINT	
<i>Protect your children</i>	non-affecting
<i>Protect yourself</i>	non-affecting
NEAR-CAVALIER CONSTRAINT	
<i>Keep your promises</i>	affecting
<i>Oppose terrorist activity</i>	affecting

Suppose now that preventing catastrophe remains above the Cavalier Constraint, which entails that it *does not affect* the permissibility of  $\phi$ . Unfortunately for the permissions to protect your children and to protect yourself, while they did fall below the Cavalier Constraint, they would (if they affected  $\phi$ ) permit an agent in treating or regarding  $\phi$  as *nearly* morally inconsequential. Thus, they fall above the Near-Cavalier Constraint, and therefore do not affect the permissibility of  $\phi$ .

A proponent of the Standard View might and should accept this scale. No reason can affect the permissibility of  $\phi$  if doing so would either permit  $\phi$  to be performed cavalierly or *nearly* cavalierly. The trouble is “nearness,” which is a paradigmatic example of a predicate that necessarily admits of vague cases. Figure 2 (above) represents protecting humanity and keeping your promises as permissions normatively effect  $\phi$ , but not in a way that would objectively permit any agent to regard doing  $\phi$  as nearly morally inconsequential. Yet among the reasons that would permit doing  $\phi$  (or would further permit doing  $\phi$ ), there are some considerations are metaphysically indeterminate with respect to affecting  $\phi$ , as represented in Figure 3.

Some considerations are weighty enough that they will nearly permit  $\phi$  to be done cavalierly, whereas others are weak enough that they will be considerably distant from doing that. Still other duties will be borderline cases: it will be neither true nor false that they *nearly* permit  $\phi$

to be done cavalierly. This occurs in ordinary cases of nearness. If two objects are within half an inch of each other, they are near each other. If they are a mile away from each other, they are not near each other. But it is indeterminate whether they are near each other if they are, say, forty-eight feet apart.

Because the permission to get answers from the terrorist is a similarly borderline case, it might be represented as *either* affecting *or* non-affecting with respect to  $\phi$ . Of course, it *can't* be the case that the permission *both affects  $\phi$  and fails to affect  $\phi$* , since the law of excluded middle holds for whether 'Get answers from the terrorist' is *permission-affective* with respect to  $\phi$ . Yet as an indeterminate case, 'Get answers from the terrorist' would do just that. Such a status is logically impossible (since excluded middle holds) and, therefore, it is impossible for such indeterminate cases to exist. But as a matter of conditional necessity, there are such cases if there is a Near-Cavalier Constraint and if any reasons whatsoever normatively affect  $\phi$ : some will affect  $\phi$ , some will not, and still others (the borderline cases) will do neither.

<u>Consideration</u>	<u>Permission-Affecting Status</u>
<i>Protect your children</i>	non-affecting
<i>Protect yourself</i>	non-affecting
<i>Get answers from the terrorist</i>	non-affecting(?)
NEAR-CAVALIER CONSTRAINT	
<i>Get answers from the terrorist</i>	affecting(?)
<i>Protect humanity</i>	affecting
<i>Keep your promises</i>	affecting

A final objection bears similarity to the last objection. There exists a view, much like the one defended here, which concedes the moral impossibility of permissibly assaulting someone in a cavalier manner. But unlike the view defended here, the alternate view denies that the Cavalier

Constraint rules out overdetermination. The view is similarly motivated: permissibly performing certain actions cavalierly appears morally impossible; therefore, it is morally impossible. If self-defensive assault is permissible at all, then overdetermined self-defensive assault is possible. But this, an objector might argue, *doesn't* rule out overdetermination, even massive or infinite overdetermination. Rather, it rules out *cavalier-making* overdetermination, including cavalier-making massive and infinite overdetermination. In effect, the alternate view claims that no matter how seriously a permission to self-defensive assault is overdetermined, it remains impossible to trivialize assaulting persons.

The alternate view's deficit concerns the implausibility of maintaining that the Cavalier Constraint is necessarily unaffected by the various relations there are, including overdetermination. The truth of the Cavalier Constraint fails to rule out as impossible any actual (or merely logically possible) relations, but it rules out certain relations that, absent the constraint, would be possible. Think of this in terms of moral theory. Some moral theories do not include the Cavalier Constraint, whereas others do. Of those that do not, certain moral relations, like *permitting-cavalier-treatment-via-overdetermination*, are maintained to be possible. Thus, the Cavalier Constraint implies that some moral theories are false.

Now consider an analogy. Suppose it is claimed that the act of rotating a ceramic penguin to the slight annoyance of a friend cannot be done in a morally casual way. Against this, it is reasonably argued that such an act can surely be done with moral ease. After all, if rotating the penguin is all that is required to prevent every bad state of affairs in the world (every war, every labored and asthmatic breath), surely some agent would be objectively justified in rotating the penguin without a second thought. The evidential basis for inferring trivializing is apparent: massive overdetermination. Said another way, the worse not- $\phi$ -ing gets, the harder it is to justify

not- $\phi$ -ing. Thus, the worse not- $\phi$ -ing, the *easier* it is to justify  $\phi$ -ing. This is evidence for the conclusion that Cavalier Constraint is sensitive to overdetermination in the following way: Massive overdetermination is *sufficient* for permitting an action to be performed in a morally cavalier way. But that is just what the defender of the alternate view denies. Thus, the alternate view is false.

## 5. Summing Up

For various reasons, most adopt and defend a particular thesis about self-defensive assault, namely:

Standard View: Some instances of self-defensive assault are morally permissible.

Here it was argued that the Standard View is false in virtue of an untoward modal implication: that there can be moral reasons to assault a person such that those reasons permit an act of self-defense, and that the reasons which permit self-defensive assault can overdetermine the permissibility for self-defensive assault. Such overdetermination is impossible for acts for which there is a Cavalier Constraint, and there is such a constraint for assaulting persons. Since the Standard View implies that there is some possible world in which assaulting persons can be permissibly trivialized and there is no such world, it is necessarily false.

Against this, it was argued that the argument from trivializing itself permits advocates of the Standard View to reject cavalier-making overdetermination while accepting overdetermination. But this is implausible because at least for the act of assaulting persons, there is also a Near-Cavalier Constraint, and *that* constraint forces the defender of the Standard View to

claim that there are metaphysically indeterminate cases of the following sort: some moral consideration, such as ‘Get answers from the terrorist’, neither *normatively permits* self-defense nor *fails to permit* self-defense. Since every moral consideration necessarily does one or the other, this is impossible, and thus there can be no moral considerations that normatively permit self-defense. Plausibly, this implies that self-defensive assault is not the sort of thing that can be permitted because it is impossible, and the reason it is impossible is that it violates at least one side-constraint on permissible action, making it absolutely wrong. Thus, the Standard View is false.

## References

Audi, Robert. (2009) “On the Meaning and Justification of Violence.” In *Violence: A Philosophical Anthology*, edited by Vittorio Bufacchi. New York: Palgrave Macmillan, 136-167.

Brummer, James. (1996) "The Structure of Residual Obligations." *Journal of Social Philosophy* 27, no. 3: 164-180.

Bufacchi, Vittorio, and Jean Maria Arrigo. (2006) "Torture, Terrorism, and the State: A Refutation of the Ticking-Bomb Argument." *Journal of Applied Philosophy* 23, no. 3: 355-373.

Davion, Victoria. (1990) "Pacifism and Care." *Hypatia* 5, no. 1: 90-100.

Doggett, Tyler. (2011) "Recent Work on the Ethics of Self-Defense." *Philosophy Compass* 6, no. 4: 220-233.

Draper, Kaila. (2009) "Defense." *Philosophical Studies* 145, no. 1 (2009): 69-88.

- . (1993) "Fairness and Self-Defense." *Social Theory and Practice* 19, no. 1: 73-92.
- Dworkin, Ronald. (1984) "Rights as Trumps." Chap. 6 in *Theories of Rights*, edited by J. Waldron. New York: Oxford University Press.
- Fiala, Andrew. (2014) "Pacifism." *Stanford Encyclopedia of Philosophy*.
- . (2014) "Pacifism and the Trolley Problem." *The Acorn* 15 (1): 33-41.
- Frowe, Helen. (2014) *Defensive Killing*. New York: Oxford University Press.
- . (2011) *The Ethics of War and Peace: An Introduction*. New York: Routledge.
- . (2010) "A Practical Account of Self-Defence." *Law and Philosophy* 29: 245-272.
- Kamm, F.M. (2012) *The Moral Target: Aiming at Right Conduct in War and Other Conflicts*. New York: Oxford University Press.
- Kellenberger, James. (1987) "A Defense of Pacifism." *Faith and Philosophy* 4 (2): 129-148.
- Lang, Gerald. (2014) "Why Not Forfeiture?" In Helen Frowe and Gerald Lang (eds.) *How We Fight*. New York: Oxford University Press, 38-61.
- Lazar, Seth. (2009) "Responsibility, Risk, and Killing in Self-Defense." *Ethics* 119, no. 4: 699-728.
- McMahan, Jeff. (2011) *Killing in War*. New York: Oxford University Press.

—. (1994a) "Revising the Doctrine of Double Effect." *Journal of Applied Philosophy* 11, no. 2: 201-212.

—. (1994b) "Self-Defense and the Problem of the Innocent Attacker." *Ethics* 104: 252-290.

—. (1994c) "Innocence, Self-Defense and Killing in War." *The Journal of Political Philosophy* 2, no. 3: 193-221.

Nagel, Thomas. (2005) "Ruthlessness in Public Life." In *Mortal Questions*, by Thomas Nagel. New York: Cambridge University Press.

Narveson, Jan. (1965) "Pacifism: A Philosophical Analysis." *Ethics* 75, no. 4: 259-271.

—. (2003) "Terrorism and Pacifism: Why We Should Condemn Both." *International Journal of Philosophy* 17, no. 2: 157-172.

Nozick, Robert. (1974) *Anarchy, State, and Utopia*. New York: Basic Books.

Orend, Brian. (2005) "War." *Stanford Encyclopedia of Philosophy*.

Otsuka, Michael. (1994) "Killing the Innocent in Self-Defense." *Philosophy & Public Affairs* 23, no. 1: 74-94.

Parfit, Derek. (2011) *On What Matters: Volume 1*. New York: Oxford University Press.

Quong, Jonathan. (2009) "Killing in Self-Defense." *Ethics* 119: 507-547.

—. (2012a) "Liability to Defensive Harm." *Philosophy & Public Affairs* 40, no. 1: 45-77.

—, and Joanna Mary Firth. (2012b) "Necessity, Moral Liability, and Defensive Harm." *Law and Philosophy* 31, no. 6: 673-701.

Rodin, David. (2002) *War and Self-Defense*. New York: Oxford University Press.

Ryan, Cheyney. (1983) "Self-Defense, Pacifism, and the Possibility of Killing." *Ethics* 93 (3): 508-524.

Steinhoff, Uwe. (2012) "Legalizing Defensive Torture." *Public Affairs Quarterly* 26: 19-32.

—. (2013) *On the Ethics of Torture*. Albany, New York: State University of New York Press.

—. (2006) "Torture -- The Case for Dirty Harry and Against Alan Dershowitz." *Journal of Applied Philosophy* 23, no. 3: 337-353.

Thomson, Judith Jarvis. (1991) "Self-Defense." *Philosophy and Public Affairs* 20, no. 4: 283-310.