

Make Them Rare or Make Them Care

Artificial Intelligence and Moral Cost-Sharing

*Blake Hereth and Nicholas G. Evans**

1. Introduction

Artificial intelligence (AI) describes a set of computational tools in which machines can evaluate, act on, and modify future behavior based on their own calculations and behaviors. AI comes in a broad range of formats, from basic personal assistants such as Apple's Siri that recognizes natural language to potential future "artificial *general* intelligence" that can make decisions over an exceptionally large range of domains and may even possess its own mental states. The military has long expressed interest in AI's processing power over very large and multimodal data sets. This power could be leveraged in areas as diverse as intelligence collection; educating soldiers or calculating precise nutrition and training regimens; and deploying AI in combat either through the use of unmanned vehicles or defending against cyberattackers or other AI.

A central component of the broader debate about military AI is lethal autonomous weapon systems (LAWS). LAWS, in brief, are armed weapon systems capable of learning, adapting, and decision-making capacities and thus in principle to make better *ethical* decisions about which targets to kill.¹ Critics such as Robert Sparrow, however, argue that moral responsibility (and attendant liability) for illegal or immoral killings by LAWS rests neither with the programmer, nor the commanding officer, nor LAWS themselves.² This, Sparrow contends, violates the *jus in bello* principle that someone can be held responsible for illegal or immoral killing.³ More recently, Michael Skerker, Duncan Purves,

* We are grateful to Tobias Vestner and Daniel Schoeni for comments on an earlier draft. Special thanks to Claire Finkelstein for inviting us to contribute to this volume.

¹ See generally RONALD C. ARKIN, *GOVERNING LETHAL BEHAVIOR IN AUTONOMOUS ROBOTS* (2009).

² See generally Robert Sparrow, *Killer Robots*, 24(1) *J. APPLIED PHIL.* 62 (2007).

³ *Id.* at 67. As Sparrow explains, "The least we owe our enemies is allowing that their lives are of sufficient worth that someone should accept responsibility for their deaths. Similarly, their grieving relatives are entitled to an answer as to why they died, which includes both knowing who is responsible and what their reasons were. Ensuring that someone can be held responsible for each death caused in war is therefore an important requirement of *jus in bello*."

and Ryan Jenkins defended a novel explanation for widespread unease about LAWS, namely, that human agents forfeit their right against harm *exclusively* to fellow humans who comprehend their value⁴ and their “functioning in response to changing circumstances in the environment in which they are deployed,” as well as being capable of making firing decisions on their own.⁵ LAWS have generated a rigorous debate in ethics.⁶ Advocates of LAWS such as Ronald Arkin will claim we can expect LAWS to eventually surpass humans in their capacity for information processing.⁷

A question arises, then, as to how we should *develop* LAWS and other military AI. According to one analysis, the market capture for military AI will exceed \$11.6 billion by 2025.⁸ The debate around LAWS and other AI applications have largely focused on when or if it is permissible to use AI, but has largely (with the exception of Sparrow) avoided the question of what kinds of AI it is permissible to *develop*. Development offers an interesting ethical quandary because, much like the pharmaceutical industry, it asks less which kinds of weapon are permissible to use and more what kinds of properties a useable weapon should have.⁹ And while there may be some kinds of weapon it is simply impermissible to develop (e.g., biological weapons), for most others the question is how we should develop LAWS and what they should be developed to do. Here we defend a novel conception of moral cost-sharing when it comes to AI and military decision-making and tie it to questions of AI development in the defense industry. Drawing from recent work on compensation in the ethics of war, we argue that the moral costs of unjust harming should be equitably distributed across responsible parties. We then argue that as some of the deepest moral costs are *psychological* in nature (e.g., the grief experienced by families of civilian “collateral damage”), justice requires that these costs be shared. In Section 2, we argue that widespread substitution of AI in critical military decisions, using LAWS as the paradigmatic example, undermines a requirement of justice in virtue of lacking moral emotions and thus being unable to share in the psychological costs of war. Then,

⁴ By this, Skerker et al. maintain that human agents can't forfeit their rights against non-agents, since that would require them to have had rights against those non-agents—a view that in turns requires non-agents to have had moral obligations, *per impossibile*. See generally Michael Skerker, Duncan Purves, & Ryan Jenkins, *Autonomous Weapons Systems and the Moral Equality of Combatants*, 22 ETHICS & INFO. TECH. 197 (2020).

⁵ Heather M. Roff, *Lethal Autonomous Weapons and Jus Ad Bellum Proportionality*, 47 CASE W. RES. J. INT'L L., 37, 37 (2015).

⁶ See, e.g., *id.*

⁷ Arkin *supra* note 1.

⁸ *Artificial Intelligence in Military Market Worth \$11.6 Billion by 2025*, MARKETSandMARKETS (Mar. 15, 2021), <https://www.prnewswire.com/news-releases/artificial-intelligence-in-military-market-worth-11-6-billion-by-2025--exclusive-report-by-marketsandmarkets-301247146.html>.

⁹ Aaron Fichtelberg, *Applying the Rules of Just War Theory to Engineers in the Arms Industry*, 12(4) SCI. & ENG'G ETHICS (2006).

in Section 3, we consider arguments from Bradley Strawser and others claiming that because LAWS enhance the safety of justified combatants, their use is morally obligatory. In response, we raise the possibility of *sentient* LAWS (SLAWS) with the capacity for moral emotion.¹⁰ Our conciliatory contention, which weds our argument against non-emotive LAWS with Strawser's argument for LAWS, is that the development of SLAWS is itself morally obligatory. Our second contention is less conciliatory: where the choice is exclusively between deploying LAWS or deploying humans in warfare, this entails a choice between prioritizing the interests of justified combatants who cause unjust harm to civilians or prioritizing the interests of the civilians whose interests they unjustly threaten. Absent a lesser-evil obligation on the part of justified combatants, justice requires prioritizing the psychological interests of civilians. We then consider what this argument might look like, abstracted to military AI in more general terms.

Unlike other arguments against LAWS, our argument doesn't conclude that the use of LAWS is all-things-considered morally impermissible to use or develop. Rather, our argument holds that the use of LAWS is pro tanto impermissible, that there are competing considerations to using LAWS (i.e., facilitating the safety of justified combatants), and that a moral dilemma emerges: either we should design LAWS capable of moral emotions, or we should limit the use of (non-emotive) LAWS. We then chart what this might look like for the defense industry to navigate this space and develop AI applications that are permissible to deploy in the use of lethal force.

2. Moral Cost-Sharing and Psychological Costs

In general, *moral cost-sharing* refers to a fair distribution of burdens between interested parties, where the moral burdens in question can be wide-ranging *financial cost-sharing*, that is, how to equitably balance economic burdens across relevant parties. For example, cost-sharing elements of the Patient Protection and Affordable Care Act caused extensive debate over the fairness of requiring healthy, responsible taxpayers to foot the bill for unhealthy, less responsible patients while often acknowledging that cost-sharing between healthy, responsible parties was equitable.¹¹ Or, to take another example, Avia Pasternak argues that citizens ought to share in the costs of injustices committed by

¹⁰ Our view is that sentience is necessary but insufficient for moral emotion. Psychopaths, for instance, are sentient, but lack moral emotion.

¹¹ See generally John P. Geyman, *Cost-Sharing Under Consumer-Driven Health Care Will Not Reform U.S. Health Care*, 40(3) J.L., MED. & ETHICS 574 (2012); Allison K. Hoffman, *Cost-Sharing Reductions, Technocrat Tinkering, and Market-Based Health Policy*, 46(4) J.L., MED. & ETHICS 873 (2018).

their state because they benefit from the intrinsically good civic bond that allows citizens to pursue, through their elected officials, morally worthy social outcomes.¹²

Ethicists and political philosophers have more recently expanded the application of moral cost-sharing to nonfinancial burdens. Sara Van Goozen, for instance, claims that insofar as noncombatants benefit from their compatriots fighting just wars that endanger foreign noncombatants, the former noncombatants ought to share in these costs.¹³ A relatively uncontroversial example of this is President Woodrow Wilson's "Meatless Mondays" campaign during the First World War, intended to limit civilian consumption of meat to conserve food for both American troops abroad and foreign civilians whose livelihoods were disrupted by global warfare. However, questions remain about the extent to which noncombatants are obligated to shoulder these costs. Had Wilson instituted "Starving Solstices" and required the American Midwest to forgo food for the duration of the winter in order to feed war-displaced foreign civilians, his humanitarian-motivated austerity would have been opposed as transparently unjust.

Against views like Van Goozen's, Helen Frowe sketches an argument intended to limit the liability of justified combatants—that is, combatants whose wartime involvement is morally justified—and, by extension, innocent (i.e., nonliable) noncombatants. Frowe's argument is meant to solve a problem raised by Jeff McMahan: Can combatants who cause collateral harm to civilians as a side effect of attacks on lawful targets be morally liable to defensive harm on that basis?¹⁴ To see the problem, imagine a case where an aerial bomber is morally justified in dropping a bomb that will (as intended) kill mostly unjust combatants—that is, combatants whose wartime involvement is morally unjustified—and merely foreseeably kill a few civilians.¹⁵ The bomber poses a justified but unjust harm to the civilians, and we typically think it is permissible to defend oneself against unjust harm. McMahan's solution is to claim that justification defeats liability,

¹² See generally Avia Pasternak, *Sharing the Costs of Political Injustices*, 10(2) POL., PHIL. & ECON. 188 (2011).

¹³ See generally Sara Van Goozen, "Sharing the Costs of Fighting Justly," 2 CRITICAL REV. INT'L SOC. & POL. PHIL. 1 (2018).

¹⁴ Persons are liable to *defensive harm* just in case they forfeit their right against others harming them in self-defense or defense of others. In such cases, persons are liable to a specific kind of instrumental harming (i.e., harm that averts an unjustified threat) that is distinct from liability to other kinds of harming, such as desert-based harming which is non-instrumental. See generally Jeff McMahan, *Self-Defense Against Justified Threateners*, in *HOW WE FIGHT: ETHICS IN WAR* 104 (Helen Frowe & Gerald Lang eds., 2014).

¹⁵ We prefer the just/unjust combatant distinction to the lawful/unlawful combatant distinction for two reasons: first, because ours is an *ethical* critique of AI-enabled LAWS that tracks primarily the morality rather than the law of war, and second, because McMahan and Frowe represent revisionist views of the Just War Theory tradition for whom the lawful/unlawful combatant distinction has no intrinsic moral importance.

such that justified threateners cannot be liable to defensive harm for the justified harms they pose.¹⁶ In short, this means that posing a *justified* threat to someone immunizes you against their justifiably harming you in self-defense.¹⁷ However, McMahan's view has been vigorously challenged.¹⁸

In clever fashion, Frowe sidesteps McMahan's solution and instead contends that whether or not lesser-evil justifications are sufficient to block moral liability, lesser-evil *obligations* are.¹⁹ Returning to the scenario, suppose the bomber was not merely justified but obligated to drop their bomb and that the obligation in question was an agent-neutral (that is, an obligation applying to all agents, irrespective of partisanship), lesser-evil (that is, an obligation to prevent the greater evil) obligation. In that case, were the civilians to kill the bomber and prevent the bombing, they would violate *their own* lesser-evil obligation to permit the bombing. As Frowe puts it:

It seems generally implausible that one might forfeit basic rights by doing one's duty, since this entails that morality might leave one with no morally permissible options for retaining one's rights. [. . .] Moreover, it seems straightforwardly incoherent that morality might pronounce Pedestrian liable to be harmed to prevent her from doing the very thing that morality requires her to do.²⁰

Accounts like Van Goozen's and Frowe's are illustrative of the more general phenomenon of moral cost-sharing. In all the aforementioned cases, cost-sharing is *zero-sum*: a finite, definite moral burden is distributed equitably.²¹ We "subtract" burdens (e.g., the civilians') from one party and "add" to another (e.g., the bomber's). Like matter itself, the costs are neither created nor destroyed, only redistributed. Call the thesis that moral cost-sharing is necessarily restricted to reallocation or redirection the *Non-Additive Thesis*.²² Having now illustrated

¹⁶ *Id.* at 118.

¹⁷ In practice, this means that insofar as the threat you pose is justified, no one can justifiably harm you to prevent you from posing that threat.

¹⁸ See generally Adam Hosein, *Are Justified Aggressors a Threat to the Rights Theory of Self-Defense?*, in *HOW WE FIGHT* 87 (Helen Frowe & Gerald Lang eds., 2014).

¹⁹ Helen Frowe, *Lesser-Evil Justifications for Harming: Why We're Required to Turn the Trolley*, 68(272) *PHIL. Q.* 460, 476–77 (2018).

²⁰ *Id.* at 476.

²¹ The standard conception of *zero-sum* is, in short, that whatever benefits one party harms the other and vice versa.

²² Our discussion of the Non-Additive Thesis resembles discussions within political philosophy of the so-called *leveling-down objection*. According to this objection, egalitarian views requiring strict equality of goods imply that justice can be satisfied merely by *removing* goods from parties with more goods than others. See generally Derek Parfit, *Equality or Priority?*, in *THE IDEAL OF EQUALITY* (Matthew Clayton & Andrew Williams eds., 2002). The Non-Additive Thesis raises a similar objection: just as it seems impermissible to remove goods from parties to achieve equality, so also it seems impermissible to *add evils* (in this case, harms) to achieve equality.

non-additive moral cost-sharing, we will defend the possibility and moral equity of *additive* moral cost-sharing.²³ Specifically, we offer a hearty defense of *solidarity-based* moral cost-sharing that, while additive in nature, explains our moral intuitions in ways the Non-Additive Thesis cannot. Said differently, we contend that moral cost-sharing concerns not simply *who suffers* or *how much someone suffers*, but also *who suffers alone*. So we will argue that the Non-Additive Thesis should be rejected.

Before examining the Non-Additive Thesis more directly, consider a powerful argument in its favor. The premises of this argument could be drawn from a number of ethical theories:

The Less Harm Argument

- P1. We should not cause nonconsensual harm unless it is practically necessary to avert a worse nonconsensual harm. [Assumption]
- P2. If P1, then the Non-Additive Thesis is true. [By definition]
- C. So the Non-Additive Thesis is true. [From P1 to P2]

The argument relies primarily on the plausibility of P1, which we hold to be extremely plausible. Once we leave room for causing consensual and lesser-evil harms, we are left with only nonconsensual harms that *do not* prevent worse harms.²⁴ Surely, it is impermissible to cause *those* harms. If so, then permissible moral cost-sharing is restricted to extant harm and prohibits the creation of new harm. So P1 by definition entails the Non-Additive Thesis, just as P 2 claims. So, as P3 follows from P1 and P2, P3 is true.

The *Less Harm Argument* fails, and with it the Non-Additive Thesis, as shown by a real-life event that serves as the central case for our essay. It's July 12, 2007, in Iraq. Crazy Horse 1-8, an American Apache helicopter, mows down 11

²³ Here's an example designed to explain the Non-Additive Thesis:

Abdul and Cleo live in a community where flashfloods are common. Abdul, who controls the emergency levies that determine where surplus water is redirected, is one day forced to decide whether to redirect flashflood waters to Cleo's home (thereby saving Abdul's home) or redirect it to his own home (thereby saving Cleo's home). If Abdul does nothing, both homes will be destroyed. There's no way to save both homes. Not wanting his home destroyed, Abdul chooses the emergency levy that preserves his home but destroys Cleo's. Later, and without Cleo ever knowing, Abdul burns his own home to the ground to alleviate his guilt.

In this example, there's something perverse about Abdul's destruction of his home. Cleo isn't helped by it, and nor is Abdul. When Abdul razes his home, he neither reallocates nor redirects any harm suffered by Cleo; rather, he merely adds new harm, which is forbidden as a form of moral cost-sharing under the Non-Additive Thesis. By contrast, imagine Abdul had another option at his disposal: to choose an emergency levy that would redirect *most* of the flood waters away from both homes but still cause *mild* damage to both homes. In the revised case, Abdul effectively minimizes the amount of harm suffered by all parties, as neither he nor Cleo must shoulder *all* the flash flood damage by having their respective homes destroyed. Thus, Abdul *reallocates* the harms posed by the flash flood—a permissible form of moral cost-sharing under the Non-Additive Thesis.

²⁴ For an example, see the case of Abdul in the previous footnote.

civilians. As the American gunner, known only as “Kyle,” fires, he exclaims, “All right! Haha! I hit ‘em.” Among the dead are two Reuters reporters, both of whom tried to escape the gunfire and were killed as they took shelter. Later, an ambulance arrived to transport the wounded to a nearby hospital. Crazy Horse 1-8 destroyed it and killed both medics. The war crimes of Crazy Horse 1-8 were initially documented in 2010 by Chris McGreal at *The Guardian*.²⁵ Footage of the crimes, including audio, was later released by Wikileaks under the title “Collateral Murder.”²⁶

This incident was indisputably a war crime. But let us examine Kyle’s emotional reaction to killing 11 civilians. As is made transparent both from the video and our reconstruction of the event, Kyle is downright *gleeful* about his lethal work. He laughs as he pumps hot lead through the bodies of his targets, demonstrating no sadness or remorse. Even without video, an attentive listener might mistakenly assume Kyle was not engaging *people* but rather simulated or inanimate targets. But he wasn’t, and they weren’t. Kyle killed human beings—people with feelings, who feared death, who suffered as they were cut down—people with families, friends, and communities who would come to miss them. And Kyle *enjoyed* cutting them down. Because he was destroying not inanimate objects but human beings, his affects were *morally ill-suited* to his behavior. Indeed, Kyle’s reaction to killing human beings—whether permissible targets or not—is morally perverse. He ought not delight in others’ suffering, death, or misfortune.

Let us assume, contrary to fact, that Kyle’s targets *were* enemy combatants and that his killing them was morally justified. Even so, Kyle should not react gleefully to gunning them down. One reason for this concerns the fact that Kyle’s targets do not exist in a moral vacuum; they are connected to others who care for them and who are adversely impacted by their demise. More relevantly, these others—partners, children, parents, friends, neighbors, imams—are not liable to the harms Kyle imposes on them. That is, they don’t forfeit their rights against Kyle that he does not harm them. Whatever they lose, they lose unjustly—even if justifiably. Like civilians harmed in tactical bombings, those who harm them act unjustly in the sense that they harm people who are not liable to be harmed. An action is *morally justified* if it is permissible to perform it, but an action is *morally just* only if all harmed parties are liable to be harmed. Indeed, civilians harmed as unavoidable side effects of military action are entitled to compensation in

²⁵ Chris McGreal, *Wikileaks Reveals Video Showing US Air Crew Shooting Down Iraqi Civilians*, THE GUARDIAN (Apr. 5, 2010), <http://www.guardian.co.uk/world/2010/apr/05/wikileaks-us-army-iraq-attack>.

²⁶ *Collateral Murder*, WIKILEAKS (Apr. 5, 2010), https://wikileaks.org/wiki/Collateral_Murder,_5_Apr_2010.

war.²⁷ Indifference toward their suffering (or, worse, *enjoyment* of it) is a moral failing, an improper response to causing others' pain.²⁸ Thus, Kyle experiences a deficiency in moral affect for the unjust harm he causes to nonliable parties with vested interests in the welfare of Kyle's victims.

We propose to explain normatively permissible moral affect in terms of cost-sharing. In cases like Kyle's, most theorists agree it is intrinsically better for *Kyle* to shoulder the costs of war than for *nonliable civilians* to suffer them, as Kyle bears greater responsibility than civilians for their endangerment. But this is often infeasible and counterproductive: if justified combatants are injured or killed, they are far less likely to achieve their justified tactical aims and thus less likely to satisfy the reasonable chance of success criterion for permissible warfare.²⁹ What, then, is the next best thing justified combatants can do? If they cannot *relieve* the suffering of nonliable parties, they can at least *share in them*. This requires, minimally, that justified combatants *care* about their innocent victims, that they avoid apathy or cruel delight. To be sure, this requirement is partially instrumental: the obligation to *act* cautiously with respect to nonliable civilian interests tends to lead to *caring* about their interests, and thus a deficiency in the latter is evidence of a deficiency in the former. Said differently, combatants who *care little* about civilians are more likely to *mistreat* them. But the obligation to care has a further, noninstrumental underpinning: we *should* care about others, *including* feeling remorseful when they suffer unjustly, *irrespective* of whether our caring about them improves their plight, and *especially* when we have wronged them.

Call this principle the *Moral Affect Principle* (MAP). So stated, MAP approaches moral bedrock: we should feel remorseful when we have wronged others.³⁰ To feel otherwise, or to feel nothing at all, divorces wrongdoing from caring and expresses to those we have wronged that we will cause them to suffer at no cost to us. In the next section, we explain the ways in which an expansive use of LAWS threatens MAP.

²⁷ See generally Saba Bazargan-Forward, *Compensation and Proportionality in War*, in *WEIGHING LIVES IN WAR* 173 (Jens David Ohlin, Larry May, & Claire Finkelstein eds., 2017); Saba Bazargan-Forward, *Defensive Wars and the Reprisal Dilemma*, 93(3) *AUSTRALASIAN J. PHIL.* 583 (2015). Cf. Jessica Flanigan & Christopher Freiman, *Drug War Reparations*, 97(2) *RES PHILOSOPHICA* 141 (2020).

²⁸ See generally Kristján Kristjánsson, *Fortunes-of-Others Emotions and Justice*, 28(105) *J. PHIL. RES.* (2003).

²⁹ See generally Ned Dobos, *Idealism, Realism, and Success in Armed Humanitarian Intervention*, 44(2) *PHILOSOPHIA* 497 (2016); Suzanne Uniacke, *Self-Defense, Just War, and a Reasonable Prospect of Success*, in *HOW WE FIGHT: ETHICS IN WAR* (Helen Frowe & Gerald Lang eds., 2014).

³⁰ See SETH LAZAR, *SPARING CIVILIANS* 108 (2015).

3. How LAWS Threaten Moral Cost-Sharing

LAWS are still some way off, and early systems will mimic sentience, mirroring human cognitive processes rather than possessing mental states.³¹ They are autonomous only in the sense that they can act without real-time direction from human operators. Autonomous drones, for instance, will run on algorithms that cause them to terminate specific targets on-sight without further human approval. Setting aside the question of whether AI weapons like LAWS might ever *become* sentient, it's clear that any near-future LAWS will be nonsentient.

Their lack of sentience, moreover, implies that they lack the full range of mental properties for which sentience is necessary. LAWS lack beliefs and other doxastic attitudes, possess neither desires nor other feelings, and are not “subjects” in the sense that there is something-it-is-like to be them.³² By implication, LAWS also lack moral affects (i.e., moral feelings) and moral responsibility generally. They are neither happy nor sad, relieved nor frustrated when they achieve their objectives. They are indifferent to human flourishing or floundering. As James Cameron's *Terminator* character Kyle Reese memorably describes Arnold Schwarzenegger's iconic Model 101 Series 800 killing machine, “It can't be bargained with. It can't be reasoned with. It doesn't feel pity, or remorse, or fear. And it absolutely will not stop until you are dead.”³³

Returning to the real-life Kyle in Iraq, imagine his entire aerial unit—pilot and gunner—had been replaced with autonomous AI, that the AI had autonomously chosen a patrol route, identified the 11 targets, and without explicit human approval eliminated them. In this counterfactual, we know at least two things. First, we know the AI did not care about the victims or those unjustly harmed by their deaths.³⁴ Their widespread replacement of human soldiers with LAWS, then, would result in far less caring—or, as we detail later, a kind of perverse concentration of care—for civilian casualties.³⁵ Second, given the argument in the previous section, we know we *should* care about others, including feeling remorseful when

³¹ See Nicholas G. Evans, *Blind Brains and Autonomous Weapon Systems*, in *LETHAL AUTONOMOUS WEAPONS* 203, 215 (Jai Galliott, Duncan MacIntosh, & Jens David Ohlin eds., 2021).

³² Cf. Daniel Stoljar, *The Semantics of “What It's Like” and the Nature of Consciousness*, 125(500) *MIND* 1161 (2016).

³³ *THE TERMINATOR* (Orion Pictures 1984).

³⁴ Tobias Vestner asks how our argument works when human soldiers don't care about the civilians they harm. Our answer is unambiguous: it strengthens our argument against LAWS, since as the number of human soldiers who care about the civilians they harm decreases (despite being obligated to care), the number of persons who should be cared for—and the strength of the obligation to care for them—increases, and with it the number of permissible LAWS (which are incapable of caring) decreases. Private correspondence on file with the authors.

³⁵ Dan Schoeni asks whether the person who switches on a LAWS might feel some remorse about the civilian casualties that will later result. Perhaps they would, but this “secondary” form of caring is less likely than the “primary” form of caring soldiers experience when they are more directly responsible for civilian casualties.

they suffer unjustly, *irrespective* of whether our caring about them improves their plight, and *especially* when we have wronged them (i.e., MAP). Thus, we have an initial argument against the use of LAWS in warfare:

The Simple MAP Argument Against LAWS

P1. We should care about others, including feeling remorseful when they suffer unjustly, irrespective of whether our caring about them improves their plight, and especially when we have wronged them. [MAP]

P2. The widespread replacement of human soldiers with LAWS would lessen the extent to which we care about others, including feeling remorseful when nonliable civilians suffer unjustly, would not improve the plight of nonliable civilians, and would constitute a wrong.³⁶

C. So the widespread replacement of human soldiers with LAWS is morally wrong. [From P1 to P2]

The Simple MAP Argument, while simple enough, contains some shortcomings. A major worry is the responsibility problem developed by Robert Sparrow:

Firstly, the possibility that the machine may attack the wrong targets may be an acknowledged limitation of the system. If the manufacturers have made this clear to those who purchase or deploy the system, then it seems they can no longer be held responsible, should this occur. . . . Secondly, and more importantly, the possibility that an autonomous system will make choices other than those predicted and encouraged by its programmers is inherent in the claim that it is autonomous. If it has sufficient autonomy that it learns from its experience and surroundings that may make decisions which reflect these as much, or more than, its initial programming. . . . At some point then, it will no longer be possible [sic] hold the programmers/designers responsible for outcomes that they could neither control nor predict.³⁷

Nor can commanding officers be held morally responsible, Sparrow argues, if “machines are really choosing their own targets.”³⁸ Nor can we hold LAWS

³⁶ Daniel Schoeni objects that widespread replacement and deployment of LAWS might redistribute human remorse to, for example, all 330 million Americans, each of whom feels .001% of the remorse that a single bomber pilot would feel if they pulled the trigger. In that case, the *aggregate* remorse would equal (if not rival) the single bomber’s remorse. We concede this possibility but reply that, in this scenario, *each person cares too little*. That is, each of us, or at least some of us, should be deeply remorseful about civilian casualties, yet no one does under the aggregate remorse scenario Schoeni describes.

³⁷ See Sparrow, *supra* note 2, at 69, 70.

³⁸ *Id.* at 71. By extension, although commanding officers might feel remorse when LAWS cause civilian casualties, the success of Sparrow’s argument (if it succeeds) entails they aren’t morally *required* to feel remorse as a result of being morally responsible for LAWS’ behavior.

themselves morally responsible, as LAWS lack moral agency.³⁹ Sparrow's arguments can be adapted to construct a new trilemma:

Sparrow's Trilemma

P1. If nonliable civilians are wronged when harmed by LAWS, they are wronged either by (a) the LAWS' programmers, (2) the LAWS' commanding officers, or (3) the LAWS themselves. [Sparrow's Assumption]

P2. Harms caused by LAWS are ones for which neither the LAWS' programmers, nor the LAWS' commanding officers, nor the LAWS themselves are morally responsible. [Sparrow's Argument]

C. So nonliable civilians are not wronged when harmed by LAWS. [From P1 to P2]

However, applied to LAWS, premise 2 is false. Sparrow concedes that when people send killer robots into war knowing the potential risks, they bear responsibility for the anticipated risks. A commonly anticipated risk of war, as it occurs in all wars, is that nonliable civilians will be harmed during the war. That's true in wars fought with or without LAWS. When commanding officers accept the risk of harming nonliable civilians by deploying LAWS, they bear responsibility for the expected unjust harm to civilians caused by LAWS. Those who bear responsibility for unjust harm are, themselves, the ones *wronging* victims of those unjust harms.

Furthermore, we have reason to reject 1: Moral responsibility is not a precondition for moral wrongdoing. Suppose I am driving around Boston and paying exceptional attention to my surroundings, local laws, and the presence (or absence) of pedestrians or other vehicles.⁴⁰ Suppose further that, despite my conscientious best intentions, I hit a motorcyclist because my brakes, which I have otherwise maintained, fail due to a manufacturing defect. Plausibly, the auto manufacturer has still *wronged* the motorcyclist despite its total lack of moral responsibility for harming them, given that both the car and the motorcyclist arguably share the road as part of a system of risk-sharing.⁴¹ Indeed, torts concerning strict liability reflect this moral principle in law: financial liability does not require willful or negligent action on behalf of a manufacturer. In auto manufacturing this is further codified in a highly regulated recall process for systemic manufacturing issues, in which the government ensures that manufacturers not only bear financial responsibility for harms done but also bears responsibility for preventing future harms that may foreseeably arise from similar causes. All

³⁹ *Id.*

⁴⁰ Because I am driving conscientiously, it follows that I am not driving negligently.

⁴¹ See generally SVEN OVE HANSSON, *THE ETHICS OF RISK: ETHICAL ANALYSIS IN AN UNCERTAIN WORLD* (2013).

that is required for this wrongdoing is that the motorcyclist has a right against which someone trespasses, and it is difficult to see why the manufacturer's ignorance undermines the motorcyclist's right that I not harm them. Here, the manufacturer's excuse would be nonculpable ignorance: they did not know the brakes would fail nor did they know they would fail in front of the motorcyclist, so the manufacturer should not be held responsible for hitting them. But others' rights against us persist despite our moral excuses. So although we wrong the motorcyclist, we are not morally responsible for doing so.

Another worry for the Simple MAP Argument is that the nonsentience of LAWS notwithstanding, human monitors are aware of LAWS-caused collateral damages and capable of moral affect. So although LAWS cannot feel remorse when they kill, their human monitors can. Maybe that is enough to rebuff the apparent impermissibility of widely deploying LAWS. However, despite recent evidence showing that remote killing *enhances* empathetic affects from controllers,⁴² the psychological disparities between *causing* remote killings and (merely) *observing* remote killings are stark.⁴³ For instance, drone operators are less prone to post-traumatic stress disorder (PTSD) than soldiers who kill face to face.⁴⁴ Our current, best evidence tells us that physical distance is inversely correlated with empathy: the farther apart people are or perceive themselves to be, the less they care.⁴⁵ Thus, a reasonable expectation of widespread deployment of LAWS, at least insofar as they replace neurotypical human soldiers, is that nonliable civilians will be cared for less.

In addition to the preceding, proponents of LAWS maintain that with sufficiently advanced weapon platforms, human soldiers may be replaced with LAWS.⁴⁶ But this would, if it came to pass, likely overburden human monitors. As the number of human soldiers declines, the burden of moral affect goes up for remote human monitors. Soldiers are no longer physically present on the battlefield or encountering civilians eye to eye, but instead are huddled inside a warehouse for remote operations. And there are fewer of them. The shift from physically present warfighters to remote operators/monitors would increase the number of remote operators/monitors. However, we could still reasonably

⁴² See generally LT. COL. WAYNE PHELPS, *ON KILLING REMOTELY: THE PSYCHOLOGY OF KILLING WITH DRONES* (2021).

⁴³ See generally Annie Brookman-Byrne, *There Is a Spectrum of Responses to Killing Far-Off Enemies*, 4(2) PSYCHOL. 40 (2020); Abraham M. Rutchick et al., *Technologically Facilitated Remoteness Increases Killing Behavior*, 73 J. EXPERIMENTAL SOC. PSYCHOL. 147 (2017).

⁴⁴ See Rutchick et al., *supra* note 43, at 147.

⁴⁵ Arianna Schiano Lomoriello et al., *Out of Sight Out of Mind: Perceived Physical Distance Between the Observer and Someone in Pain Shapes Observer's Neural Empathic Reactions*, 9 FRONTIERS PSYCHOL. 1824 (2018).

⁴⁶ See generally Arkin, *supra* note 1. Cf. Evans, *supra* note 30; Nicholas Greig Evans, *Emerging Military Technologies: A Case Study in Neurowarfare*, in *NEW WARS AND NEW SOLDIERS: MILITARY ETHICS IN THE CONTEMPORARY WORLD* 105 (Paolo Tripodi & Jessica Wolfendale eds., 2011).

expect a considerably smaller number of soldiers bearing roughly the same psychological burden initially borne by hundreds of thousands more.⁴⁷ This is not unique to LAWS—famously, Joanna Bourke’s study of killing in war noted that artillerymen experience what is arguably a form of moral cost-sharing through their concern and suffering at distant casualties they cannot see⁴⁸—but given advances in human-machine interfaces and the possibility of controlling multiple LAWS simultaneously,⁴⁹ there is reason to fear that the toll may become extraordinary.

4. The Cost-Sharing Dilemma

MAP *prima facie* rejects the use of LAWS. After all, without the ability to engage in plausible or defensible moral cost-sharing, it seems that there is—another—pro tanto reason to abandon the use of “killer robots” in favor of humans. This would then place us in conflict with, for example, Ronald Arkin and Bradley Strawser who have both argued for a strong moral duty to deploy LAWS.⁵⁰

But things need not be that simple. Nothing in our argument rides on LAWS, and indeed in the following we extend the argument to moral cost-sharing in other AI defense systems. Rather, the argument here is only against LAWS that are incapable of engaging in moral cost-sharing. All near-future LAWS fall into this category, by virtue of lacking mental states.

But it is plausible that future LAWS could be developed to have some kind of sentience, and thus be in principle capable of experiencing moral sympathies in a way that allows for MAP to be satisfied. Over the last 10 years significant advances have been made that suggest that while current AI applications may not be sentient, there may be some artificial *general* intelligence (AGI) that could possess the capacity for mental states. This may be in virtue of sufficiently broad and deep computing power to generate mental states or the capacity for sentience through novel computing methods that allow computers to perform abductive and other human-like cognition.⁵¹ While disagreements remain as to

⁴⁷ Tobias Vestner makes the critical observation that this concern also holds for other groups, including aerial bombers whose targets are invisible to the naked eye. Fortunately, our argument against the use of LAWS to replace human soldiers doesn’t require that LAWS be the *only* technology indicted. Correspondence on file with the authors.

⁴⁸ See generally JOANNA BOURKE, *AN INTIMATE HISTORY OF KILLING: FACE TO FACE KILLING IN TWENTIETH CENTURY WARFARE* (2000).

⁴⁹ See NICHOLAS G. EVANS, *THE ETHICS OF NEUROSCIENCE AND NATIONAL SECURITY* 72–85 (2021).

⁵⁰ See generally ARKIN, *supra* note 1; Bradley Jay Strawser, *Moral Predators: The Duty to Employ Uninhabited Aerial Vehicles*, 9(4) *J. MIL. ETHICS* 342 (2010).

⁵¹ See generally ERIK J. LARSON, *THE MYTH OF ARTIFICIAL INTELLIGENCE: WHY COMPUTERS CAN’T THINK THE WAY WE DO* (2021). We note there are some difficulties here about knowing

exactly what form moral agency for machines might take, there is general agreement that if sentience is possible, moral status is also possible.⁵² For those skeptical of the likelihood or possibility of sentient LAWS, our position (developed below) is that nonsentient LAWS ought not replace human soldiers.

If proponents such as Arkin and Strawser are to be believed *and* LAWS can be designed that satisfy the MAP, then they would be obligatory to design and deploy. That is, agents that are capable of executing the acts of armed conflict with greater attention to avoiding noncombatant deaths *and* are capable in engaging in moral cost-sharing are in principle better than human agents who fail to achieve the former but satisfy the latter. So here we can act in a conciliatory manner toward advocates for LAWS, by claiming that LAWS are obligatory—if they are sentient. It may be that opponents of AI engaged in lethal decision-making retain some simple, emotive concern to AI killing—but in this case, so do the AI applications!

This possibility raises a number of interesting empirical questions that we sketch but set aside here. There will first be questions of what kinds of sentience, and how much, will be required to engage in MAP. Surely human levels of sentience are excessive for this purpose—Corvids and other animals display sufficient levels of sentience to allow for this. The precise architecture of sentient LAWS (SLAWS) will also have to be determined, and what counts as sympathies will need to be understood to a much higher degree to *know* or have appropriate confidence that sentience is in fact achieved.⁵³

The normative questions, on the other hand, are deeper. The first is whether creating sentient machines *specifically* to place them in harm's way, of the kind we describe previously, is morally permissible.⁵⁴ An objection here might be that there is a perversity to creating a sentient creature, as far as an AI application is a creature, solely so that we can conscript it to undertake our risky aims. But creating SLAWS does just that. The argument against this is simple: if we believe (with Immanuel Kant) that we should not treat self-governing agents as a mere means (since they are self-governing, rational agents with the capacity and right to self-determination), and if the sole reason we are creating a self-governing, sentient agent is to deploy it in combat and thereby expose it to risks to which it neither consented nor is liable, then we have treated that agent as a mere means.

whether AGI is “really” sentient, but point out that this is, on a number of accounts of philosophy of mind, also a challenge with other humans.

⁵² Steve Torrance, *Machine Ethics and the Idea of a More-Than-Human Moral World*, in *MACHINE ETHICS* 115 (Michael Anderson & Susan Leigh Anderson eds., 2011).

⁵³ See EVANS, *supra* note 49, at 37–47.

⁵⁴ See Alexander A. Guerrero, *Appropriately Using People Merely as a Means*, *CRIM. L. & PHIL.* 777, 785 (2016).

Worse, the *way* in which we have treated them as a mere means is exposing them to unjust harm—a practice that hardly seems justified.

As an analogy, consider recent work on the ethics of using sex robots. On the view defended by Lily Frank and Sven Nyholm,⁵⁵ there are moral reasons favoring consent-capable sex robots. Sparrow, by contrast, holds that consent-capable sex robots carries a risk of enabling rape fantasies: if sex robots can give consent, they can also withhold it, and if they can withhold it, then they can be (at least expressively) raped.⁵⁶ Imagine a sentient sex robot created for the express purpose of sex with humans. Absent an opportunity to withhold consent, humans using those robots sexually would be *sexually assaulting* them. By analogy, humans using SLAWS as mere means, where those same SLAWS are then destroyed in combat, are liable to a similar moral charge: murder.

A more compelling concern, and one in which we are less conciliatory, is what we do in the meantime. SLAWS are not yet possible. It is not clear they will *ever* be possible, as experts disagree about whether sentient robots are in principle achievable, or in practice achievable given the current trajectory of modern computer science. So, for now (and potentially forever), our choice is between deploying LAWS and humans in warfare.⁵⁷

What this entails is a choice between prioritizing the interests of justified combatants who cause unjust harm to civilians, on the one hand, and prioritizing the interests of civilians whose interests they unjustly threaten, on the other. For the reasons outlined earlier, most would favor the former disjunct of this set of choices. This is because while just combatants are permitted to engage in acts of lethal force, they are not permitted to wrong noncombatants unjustly. This, we have argued, includes the psychological costs of war.

We concede that there may be a lesser-evil justification for *not* prioritizing the interests of civilian psychological integrity over justified combatant psychological integrity. For instance, we can imagine a case wherein severe psychological harm to justified combatants makes them less likely to achieve their justified aims, such as the humanitarian rescue of endangered civilians. For a more detailed case, imagine that justified combatants have placed enormous side-effect burdens on the civilian population up to a certain point in the war and, if the justified combatants do not preserve their own mental health at the cost of additional

⁵⁵ See generally Lily Frank & Sven Nyholm, *Robot Sex and Consent: Is Consent to Sex Between a Robot and a Human Conceivable, Possible, and Desirable?*, 25(3) ARTIFICIAL INTELLIGENCE & L. 305 (2017).

⁵⁶ By “expressively,” we mean only that the nonconsensual use of robots for sex *looks* like rape and risks reinforcing the immoral message that consent is unnecessary for permissible sex. See generally Sparrow, *supra* note 1.

⁵⁷ A third option, suggested to us by Daniel Schoeni, is the creation of LAWS that *behave as if* they have moral affect. However, our position is that this is morally unsatisfactory for the same reason it’s unsatisfactory for soldiers merely to behave as if they care about civilian casualties, namely, because *actually caring* is what’s required by MAP.

civilians' mental health, then past side-effect harms suffered by civilians will have been in vain. In such a case, what McMahan calls "wide proportionality" requires that the goods secured by prior civilian suffering outweigh that prior suffering.⁵⁸ Where the securement of such goods is threatened by declining mental health among justified combatants, it follows that justified combatants ought to prioritize their mental health over others' (including civilians') mental health to satisfy wide proportionality.

We might, finally, extend this to the development and deployment of other AI systems that replace human functions in ways that ought to be meaningfully cost-shared. Consider the analyst who sifts through intelligence signals in search of a target. When they find that potential target, there is often a significant margin of error around whether that target is (1) the person the analyst thinks they are, and (2) occupying the role in a military or terrorist group they are said to occupy. Even if they are right in whom they target, however, we might think they ought to engage in the moral cost-sharing involved in the knowledge that the person targeted is a sibling, parent, spouse, and so on. They may also live knowing they never really know, *for sure*, that they killed the right person.

The trend of AI-substituted judgments for human actors has been a long time coming. AI has been suggested, since the mid-90s, as a replacement for these kinds of analytic processes. This effort began with the Total Information Awareness program that preceded 2001, became infamous immediately after it, and was shut down—but whose central functions were taken up by the National Security Agency in what would become the U.S. mass surveillance apparatus revealed by Edward Snowden.⁵⁸

All these AI enterprises fall victim to analogous concerns as LAWS in terms of MAP. Through these innovations, the defense industry is engaged in a process of rearranging, through the evolution of technologies, national security decision-making, and moral cost-sharing. The process by which an individual is targeted and eliminated by a state is now rarely an individual process of defensive action in a conventional setting of armed conflict. It is an institutional decision, made through a collaboration of discrete actors who determine separately which targets exist, which ought to be subject to lethal force, the timing and structure of that force, its authorization, and its ultimate execution. The military calls this a "kill chain." By developing novel AI, and especially broad general intelligences that are capable of synthesizing complex information and making decisions on the use of lethal force, the defense industry is engaged in—indeed, may always have been—an active, knowing, and responsible participant in the kill chain.⁵⁹

⁵⁸ See JEFF MCMAHAN, *KILLING IN WAR* 21 (2009). Some harmful act is proportionate in the wide sense just in case it doesn't cause more harm to nonliable parties than it prevents.

⁵⁹ Evans, *supra* note 49, at 11–24. We concur with Helen Frowe that agents (individual or group) are morally liable in proportion of the size of the threat to which they contribute and *not* in proportion

Given this, the moral cost-sharing of AI automation of lethal action in armed conflict should be treated with the same care as we suggest earlier. Incorporating sentience would allow AI to engage in obligatory moral cost-sharing and would relieve individuals from becoming increasingly focused sites of (justified) psychological load in the decision to kill others.

We suspect some will find this conclusion repugnant. After all, we are positing that “basic” AI is not sufficient, and instead are suggesting that we *intentionally* push further into the AI that is the stuff of science fiction nightmares. Why would we forgo ED-209 of *RoboCop*⁶⁰ and jump straight to the Skynet of *Terminator*? And why does that sound so horrific?

This is partly an empirical question, but we want to suggest one possible answer. What makes LAWS and other AI weapons so attractive to military decision makers and proponents is *precisely* that it portrays a cooler, less morally costly form of war. Christian Enemark, who notes this is a key attraction of drones and LAWS in turn, writes that “political leaders, having less cause to contemplate the prospect of deaths, injuries and grieving families, might accordingly feel less anxious about the using force to solve political problems.”⁶¹ While Enemark was primarily referring to *allied* casualties, we think this likely applies to our feelings of enemy or noncombatant casualties as well. The defense industry can market the use of sentient AI precisely because it obviates our need for moral cost-sharing. It relieves us of a burden in weighty moral decision-making and renders it impersonal and algorithmic.

Our argument is quite simply that the moral weight of these decisions ought not be obviated. We do not need to deny that AI may make certain kinds of acts safer or less costly in terms of human lives. But we note that obviating moral cost-sharing is impermissible in the use of AI for lethal force. And if the defense industry will not stop making AI applications that make these decisions, then they ought to make AI applications that are capable of bearing that cost-sharing with us, rather than creating a sink in which we can hide the moral weight of, to borrow a phrase from Michael Walzer, the “crime of war.” Moreover, insofar as the defense industry bears causal and moral responsibility for its position in the “kill chain,” they too must share some of the costs of war.

to the size of their contribution to a threat. So the defense industry (*qua* group-agent) is liable in proportion to the size of the SLAWS threat to which it contributes, which is potentially significant. See generally HELEN FROWE, DEFENSIVE KILLING 175–77 (2014).

⁶⁰ ROBOCOP (Orion Pictures 1987).

⁶¹ CHRISTIAN ENEMARK, ARMED DRONES AND THE ETHICS OF WAR: MILITARY VIRTUE IN A POST-HEROIC AGE 22 (2015).

6. Conclusion

Interest in the use of AI in warfare has increased substantially over the last 20 years and continues to expand in quantity and application. Our chapter raised a novel objection to the implementation of autonomous weapons, namely, that their replacement of human soldiers severely curtails moral cost-sharing. The debate about the use of AI in defense typically focuses on the accuracy with which those AI applications can inflict harm on combatants and spare civilians. However, as we noted at the outset, civilians are never *spared* in war. At the absolute minimum, no one belongs to no one,⁶² and all persons—even those justly killed—are part of a web of social support that means that when a combatant is killed, civilians are wronged. Part of the process of war is the moral cost-sharing of this burden for killing. AI subverts this burden in unjust ways.

To grasp the basics of our argument, consider the case of uninhabited aerial vehicles that act autonomously (i.e., LAWS). Imagine that a LAWS terminates a military target and that five civilians die as a side effect of the LAWS bombing. Because LAWS lack *moral* agency, and in particular the capacity for moral emotions, moral cost-sharing is limited to dead civilians and their loved ones. We argue that's unjust insofar as those responsible for unjust harm to others *ought* to share in those costs. Our worry expands to other strategic uses of AI, for example, cyber warfare. Here we contribute to this debate by outlining the Moral Affect Principle and putting it through its paces. Along the way, we provide a way for proponents of military AI to have their way: simply make AI sentient and capable of moral affect. However, we note that this is a bridge too far even for proponents of military AI, who increasingly are reticent to admit to this developmental and technological possibility.

This presents a dilemma: either we design autonomous weaponry capable of moral emotions, or we limit the use of autonomous weaponry. We then consider the more general case for any AI. The former undermines the risk-mitigation purpose of creating autonomous weaponry and expands the number of sentient individuals whose welfare is risked in war. The latter risks worsening combatant casualties and achieving strategic aims.

References

- RONALD C. ARKIN, *GOVERNING LETHAL BEHAVIOR IN AUTONOMOUS ROBOTS* (2009).
 Saba Bazargan-Forward, *Defensive Wars and the Reprisal Dilemma*, 93(3) AUSTRALASIAN J. PHIL. 583 (2015).

⁶² See generally EVA FEDER KITTAY, *LOVE'S LABOR: ESSAYS ON WOMEN, EQUALITY, AND DEPENDENCY* (1999).

- Saba Bazargan-Forward, *Compensation and Proportionality in War*, in *WEIGHING LIVES IN WAR* 173 (Jens David Ohlin, Larry May, & Claire Finkelstein eds., 2017).
- JOANNA BOURKE, *AN INTIMATE HISTORY OF KILLING: FACE TO FACE KILLING IN TWENTIETH CENTURY WARFARE* (2000).
- Annie Brookman-Byrne, *There Is a Spectrum of Responses to Killing Far-Off Enemies*, 4(2) *PSYCHOL.* 40 (2020).
- Ned Dobos, *Idealism, Realism, and Success in Armed Humanitarian Intervention*, 44(2) *PHILOSOPHIA* 497 (2016).
- CHRISTIAN ENEMARK, *ARMED DRONES AND THE ETHICS OF WAR: MILITARY VIRTUE IN A POST-HEROIC AGE* (2015).
- Nicholas Greig Evans, *Emerging Military Technologies: A Case Study in Neurowarfare*, in *NEW WARS AND NEW SOLDIERS: MILITARY ETHICS IN THE CONTEMPORARY WORLD* 105 (Paolo Tripodi & Jessica Wolfendale eds., 2011).
- NICHOLAS G. EVANS, *THE ETHICS OF NEUROSCIENCE AND NATIONAL SECURITY* (2021).
- Nicholas G. Evans, *Blind Brains and Moral Machines: Neuroscience and Autonomous Weapon Systems*, in *LETHAL AUTONOMOUS WEAPONS: RE-EXAMINING THE LAW AND ETHICS OF ROBOTIC WARFARE* 2013 (Duncan MacIntosh, Jai Galliot, & Jens David Ohlin eds., 2021).
- Nicholas G. Evans, *Blind Brains and Moral Machines: Neuroscience and Autonomous Weapon Systems*, in *LETHAL AUTONOMOUS WEAPONS* 203 (Jai Galliot, Duncan MacIntosh, & Jens David Ohlin eds., 2021).
- Aaron Fichtelberg, *Applying the Rules of Just War Theory to Engineers in the Arms Industry*, 12(4) *SCI. & ENG'G ETHICS* 685–700 (2006).
- Jessica Flanigan & Christopher Freiman, *Drug War Reparations*, 97(2) *RES PHILOSOPHICA* 141 (2020).
- Lily Frank & Sven Nyholm, *Robot Sex and Consent: Is Consent to Sex Between a Robot and a Human Conceivable, Possible, and Desirable?*, 25(3) *ARTIFICIAL INTELLIGENCE & L.* 305 (2017).
- Helen Frowe, *Lesser-Evil Justifications for Harming: Why We're Required to Turn the Trolley*, 68(272) *PHIL. Q.* 460 (2018).
- HELEN FROWE, *DEFENSIVE KILLING* (2014).
- John P. Geyman, *Cost-Sharing Under Consumer-Driven Health Care Will Not Reform U.S. Health Care* 40(3) *J.L., MED. & ETHICS* 574 (2012).
- Alexander A. Guerrero, *Appropriately Using People Merely as a Means*, *CRIM. L. & PHIL.* 777 (2016).
- SVEN OVE HANSSON, *THE ETHICS OF RISK: ETHICAL ANALYSIS IN AN UNCERTAIN WORLD* (2013).
- Allison K. Hoffman, *Cost-Sharing Reductions, Technocrat Tinkering, and Market-Based Health Policy*, 46(4) *J.L., MED. & ETHICS* 873 (2018).
- Adam Hosein, *Are Justified Aggressors a Threat to the Rights Theory of Self-Defense?*, in *HOW WE FIGHT* 87 (Helen Frowe & Gerald Lang eds., 2014).
- EVA FEDER KITTAÏ, *LOVE'S LABOR: ESSAYS ON WOMEN, EQUALITY, AND DEPENDENCY* (1999).
- Kristján Kristjánsson, *Fortunes-of-Others Emotions and Justice*, 28(105) *J. PHIL. RES.* 105–128 (2003).
- ERIK J. LARSON, *THE MYTH OF ARTIFICIAL INTELLIGENCE: WHY COMPUTERS CAN'T THINK THE WAY WE DO* (2021).
- SETH LAZAR, *SPARING CIVILIANS* (2015).

- Arianna Schiano Lomoriello et al., *Out of Sight Out of Mind: Perceived Physical Distance Between the Observer and Someone in Pain Shapes Observer's Neural Empathic Reactions*, 9 FRONTIERS PSYCHOL. 1824 (2018).
- MARKETSANDMARKETS, *Artificial Intelligence in Military Market Worth \$11.6 Billion by 2025* (Mar. 15, 2021), <https://www.prnewswire.com/news-releases/artificial-intelligence-in-military-market-worth-11-6-billion-by-2025--exclusive-report-by-market-sandmarkets-301247146.html>.
- Chris McGreal, *Wikileaks Reveals Video Showing US Air Crew Shooting Down Iraqi Civilians*, THE GUARDIAN (Apr. 5, 2010), <http://www.guardian.co.uk/world/2010/apr/05/wikileaks-us-army-iraq-attack>.
- JEFF McMAHAN, KILLING IN WAR (2009).
- Jeff McMahan, *Self-Defense Against Justified Threateners*, in HOW WE FIGHT: ETHICS IN WAR 104 (Helen Frowe & Gerald Lang eds., 2014).
- Derek Parfit, *Equality or Priority?*, in THE IDEAL OF EQUALITY (Matthew Clayton & Andrew Williams eds., 2002).
- Avia Pasternak, *Sharing the Costs of Political Injustices*, 10(2) POL., PHIL. & ECON. 188 (2011).
- LT. COL. WAYNE PHELPS, ON KILLING REMOTELY: THE PSYCHOLOGY OF KILLING WITH DRONES (2021).
- ROBOCOP (Orion Picture 1987).
- Heather M. Roff, *Lethal Autonomous Weapons and Jus Ad Bellum Proportionality*, 47 CASE W. RES. J. INT'L L. 37 (2015).
- Abraham M. Rutchick et al., *Technologically Facilitated Remoteness Increases Killing Behavior*, 73 J. EXPERIMENTAL SOC. PSYCHOL. 147 (2017).
- Michael Skerker, Duncan Purves, & Ryan Jenkins, *Autonomous Weapons Systems and the Moral Equality of Combatants*, 22 ETHICS & INFO. TECH. 197 (2020).
- Robert Sparrow, *Killer Robots*, 24(1) J. APPLIED PHIL. 62 (2007).
- Daniel Stoljar, *The Semantics of "What It's Like" and the Nature of Consciousness*, 125(500) MIND 1161 (2016).
- Bradley Jay Strawser, *Moral Predators: The Duty to Employ Uninhabited Aerial Vehicles*, 9(4) J. MIL. ETHICS 342 (2010).
- THE TERMINATOR (Orion Picture 1987).
- Steve Torrance, *Machine Ethics and the Idea of a More-Than-Human Moral World*, in MACHINE ETHICS 115 (Michael Anderson & Susan Leigh Anderson eds., 2011).
- Suzanne Uniacke, *Self-Defense, Just War, and a Reasonable Prospect of Success*, in HOW WE FIGHT: ETHICS IN WAR (Helen Frowe & Gerald Lang eds., 2014).
- U.S. DEP'T OF DEF., *Unmanned Aircraft Systems (UAS): DoD Purpose and Operational Use*, <https://dod.defense.gov/UAS/>.
- Sara Van Goozen, "Sharing the Costs of Fighting Justly," 2 CRITICAL REV. INT'L SOC. & POL. PHIL. 1 (2018).
- WIKILEAKS, *Collateral Murder* (Apr. 5, 2010), https://wikileaks.org/wiki/Collateral_Murder,_5_Apr_2010.