

Autonomous Weapons Systems and the Moral Equality of Combatants

Accepted to *Ethics and Information Technology*

Michael Skerker, Duncan Purves, and Ryan Jenkins

Abstract (206 words)

To many, the idea of autonomous weapons systems (AWS) killing human beings is grotesque. Yet critics have had difficulty explaining why it should make a significant moral difference if a human combatant is killed by an AWS as opposed to being killed by a human combatant. The purpose of this paper is to explore the roots of various deontological concerns with AWS and to consider whether these concerns are distinct from any concerns that also apply to long-distance, human-guided weaponry. We suggest that at least one major driver of the intuitive moral aversion to lethal AWS is that their use disrespects their human targets by violating the martial contract between human combatants. On our understanding of this doctrine, service personnel cede a right not to be directly targeted with lethal violence to other human agents alone. Artificial agents, of which AWS are one example, cannot understand the value of human life. A human combatant cannot transfer his privileges of targeting enemy combatants to a robot. Therefore, the human duty-holder who deploys AWS breaches the martial contract between human combatants and disrespects the targeted combatants. We consider whether this novel deontological objection to AWS forms the foundation of several other popular yet imperfect deontological objections to AWS.

Scholars have voiced objection to the development and deployment of fully autonomous weapon systems (AWS) largely for

reasons having to do with technical limitations or political ramifications. For example, some have raised concerns that such systems might make political leaders more cavalier about engaging in military action;¹ that non-state actors would readily be able get ahold of small and relatively cheap AWS;² or that such systems would be vulnerable to hacking. Still other scholars raise principled objections to the use of AWS, suggesting that such systems should not be deployed even if the political and technical concerns could be addressed or shown to be no different in kind than concerns about widely accepted long-distance weaponry. Some have worried that AWS would not be able to adhere to principles of Distinction and Proportionality and would thereby violate the Laws of Armed Conflict or *jus in bello* restrictions;³ that there would be no accountability for war crimes perpetrated by robots;⁴ and that the use of AWS demonstrates a “profound disrespect” for the enemy.⁵

To many, the idea of robots autonomously deciding to kill human beings is grotesque, and yet critics have had difficulty explaining why it should make a significant moral difference if a human combatant is killed by an AWS as opposed to being killed by a human combatant.⁶ It has been especially challenging to articulate the moral difference between AWS and widely accepted and commonly used forms of long-distance human-controlled weapons. Some have suggested that concerns with AWS such as the epistemic uncertainty of the authorizing human agent, the metaphysical indeterminacy of the AWS's victims, and the lack of an interpersonal relationship between human agent and victim, are also relevant to long-distance, human controlled weapons. If this is the case, the conventional acceptability of long-distance, human

¹ Docherty (2012). Docherty raises a number of additional objections to the deployment of AWS including some of the following ones on this list. Also see Kahn (2017).

² Burri (2017).

³ Armin Krishnan (2009); Guarini M, Bello P. (2012); Sharkey (2007: 122).

⁴ Sparrow (2007); Roff (2013).

⁵ Sparrow (2016a).

⁶ Sparrow cf. Robillard (forthcoming); Burri, (2017).

controlled weapons would seem to create grounds for accepting the use of AWS.⁷ The purpose of this paper is to explore the roots of various deontological concerns with AWS—concerns that seek to identify a principled *constraint* against the use of AWS—and to consider whether these concerns are distinct from any concerns that also apply to long-distance, human-guided weaponry. We will not take a stand as to whether deontological concerns trump consequentialist considerations in favor of using or not using AWS in future conflicts. We hope our discussion contributes to a resolution of this debate. We suggest that at least one major driver of the intuitive moral aversion to lethal AWS is that their use *disrespects* their human targets by violating the martial contract between human combatants. On our understanding of this doctrine, service personnel cede a right not to be directly targeted with lethal violence to other human agents alone. This ceding is not to humans *per se*, or humans by fiat, but to agents able to reason about moral considerations (including whether to exercise their rights at others' expense). Such agents would also need to have an understanding of the authority of moral reasons, reasons grounded in the value of human life. Artificial agents, of which AWS are one example, cannot understand the value of human life because they lack the experience of having personal projects or sensing their own mortality. Since a duty-holder cannot transfer his duties and privileges to an entity incapable of bearing those duties and privileges, a human combatant cannot transfer his privileges of targeting enemy combatants to a robot. Therefore, the human duty-holder who deploys AWS breaches the martial contract between human combatants and thereby disrespects the targeted combatants.

1 Some deontological objections to AWS

The most interesting objections to AWS are deontological in flavor. In this section, we catalogue what we take to be the major deontological objections to the development and deployment of AWS and identify their weaknesses. In the next section, we develop

⁷ [Blinded for peer review]

a novel deontological objection to the development and deployment of AWS.

To begin our brief survey of deontological objections to AWS it is important to be clear about what we take to be conditions under which an objection counts as a *deontological* objection to AWS. Generally speaking, deontological moral theories impose *constraints* on the promotion of the good.⁸ That is, whatever we take to determine the intrinsic value of states of affairs, a deontological moral theory will sometimes prohibit the maximal promotion of that value.⁹ It will require that we sometimes do less than the best. Therefore, we take a ‘deontological’ objection to AWS to be any objection that would count against the use of AWS *even if* AWS were to yield optimal outcomes vis-à-vis our legitimate military aims. Suppose that our legitimate military aims are achieving our just cause while minimizing casualties to civilians, enemy combatants, and our own military personnel.¹⁰ A deontological objection to AWS would morally prohibit the deployment of AWS in cases where AWS would achieve just cause while minimizing casualties to civilians, enemy combatants, and our own military personnel.

Because we are interested in deontological objections to AWS, we can set aside some familiar concerns. These are technical concerns whose force depends on the current state of artificial intelligence (AI) and which would thus lose their force were AI to reach a sufficient level of sophistication. For example, some have

⁸ We do not mean to suggest that a theory or objection *must* posit the existence of constraints to count as deontological. See Scheffler (1984) for one exception. *Generally speaking*, however, deontological moral theories posit constraints.

⁹ McNaughton, David, and Piers Rawling (1998).

¹⁰ This is admittedly an idiosyncratic way of characterizing the distinction between consequentialism and deontology and their relationship to Just War Theory, since some have argued that the principles of discrimination and proportionality are grounded in non-consequentialist moral constraints (e.g., Nagel 1972). Still, for the purposes of our discussion, it is helpful to characterize discrimination and proportionality as valuable goals to be achieved in wartime rather than as constraints on the achievement of our goals.

worried that AWS would be ill-suited to discriminate between legitimate and illegitimate targets in chaotic environments.¹¹ The inability to distinguish between legitimate and illegitimate targets leads to a number of worries about whether AWS could satisfy *jus in bello* principles of discrimination and proportionality.¹² These obstacles should not be underestimated, but they might be obviated in the near future by limiting the domain in which AWS operate,¹³ and overcome in the distant future if AI technology becomes sufficiently sophisticated.¹⁴ If AWS were to become *perfect* replacements for human soldiers, we could use them to achieve the just cause of the war without the threat of indiscriminate or disproportionate harm. But not all objections to AWS can be resolved simply by advancing the state of the art. These are the deontological objections with which we are concerned.

'Responsibility Gaps' and Respect

The natural way to begin a survey of deontological objections to AWS is to imagine a future in which AI has reached a level of sophistication that rivals that of human agents. Robert Sparrow, in his influential 2007 article “Killer Robots,” imagines a state of AI where robot agency is, in virtually all functional respects, just like human agency.¹⁵ The root of Sparrow’s worry is that, under these circumstances, there would be no one we could appropriately

¹¹ Guarini and Bello (2012); Schmitt and Turner (2013). See also Asaro, Peter. "On banning autonomous weapon systems: human rights, automation, and the dehumanization of lethal decision-making." *International Review of the Red Cross* 94.886 (2012): 687-709.

¹² For an excellent extended discussion of these challenges, see Sparrow (2015 and 2016a).

¹³ See Schmitt (2013), Guarini and Bello (2012) and Sparrow (2016a).

¹⁴ See Bostrom (2014) for a description of some of the paths to general AI that is far superior to human general intelligence.

¹⁵ It is crucial to his argument, however, that robot agency falls short of full-blooded moral agency. The origins of this ‘responsibility gaps’ objection to AWS can be found in Matthias (2004). See also Roff (2013) for an extended discussion of responsibility and liability for AWS.

hold responsible for the decisions of AWS. Sparrow offers the following argument against AWS:

1. It is morally impermissible to wage war unless we are able to justly hold someone morally responsible for the deaths of enemy combatants that we cause.
2. Neither the programmer of AWS nor its commanding officer could justly be held morally responsible for the deaths of enemy combatants caused by AWS.
3. We could not justly hold AWS morally responsible for its actions, including actions that would be war crimes had they been committed by a human soldier.
4. There are no other plausible candidates whom we might hold morally responsible for the deaths of enemy combatants caused by AWS.
5. From 2, 3, and 4: Therefore, there is no one whom we may justly hold responsible for the deaths of enemy combatants caused by AWS.
6. From 1 and 5: Therefore, it is impermissible to wage war through the use of AWS. To do so would be to treat our enemy like vermin, as though they may be exterminated without moral regard at all.¹⁶

Elsewhere, two of the authors of this paper have suggested that Sparrow's argument would lose its force if AWS ever became flawless replacements for human soldiers. If AWS never makes a mistake in warfare, it is mysterious why it would be problematic that, *were it to make a mistake*, there *would be* no one to hold morally accountable.¹⁷ Recently, however, Sparrow has further

¹⁶This summarized version of Sparrow's argument is taken from [blinded for peer review].

¹⁷ [Blinded for peer review.]

elaborated that using highly sophisticated AWS to wage war “treats our enemy like vermin” because it fails to establish an “interpersonal relationship” with the enemy that justly waging war requires. Drawing on Thomas Nagel’s canonical “War and Massacre,” Sparrow argues that waging war justly requires “establishing this interpersonal relationship with those who are the targets of a lethal attack... and acknowledging the morally relevant features that render them combatants or otherwise legitimately subjected to a risk of being killed.”¹⁸ But, Sparrow claims, this interpersonal relationship cannot be established when an AWS decides to target someone, because even a highly sophisticated AWS is not a full-blooded moral agent capable of appreciating or judging its reasons for action. It is not a member of the Kantian “Kingdom of Ends.” For this reason, “in some fundamental sense, there is no one who decides whether the target of the attack should live or die... the agency of the robot would intervene so as to make it implausible to describe the person who launched the robot as having killed anyone at all.”¹⁹ And this, Sparrow suggests, is “profoundly disrespectful” to the enemy.²⁰

But, as Sparrow acknowledges,²¹ it is questionable whether an AWS can *sever* the interpersonal relationship between the target of the AWS and the person who deploys the AWS unless AWS are full-blooded moral agents. And in that case, the responsibility gap dissolves, because AWS *would be* members of the Kingdom of Ends, capable of acknowledging the morally relevant features of combatants that render them liable to harm.

Furthermore, two of the authors of this paper have argued that it is consistent with the considerations advanced by Nagel in “War and Massacre” that AWS could be *more* respectful of combatants and non-combatants than conventional means of warfare.²² First, if AWS become vastly superior to human soldiers at discrimination between legitimate and illegitimate targets, then

¹⁸ Sparrow (2016a: 106).

¹⁹ Sparrow (2016b: 402).

²⁰ Sparrow (2016a: 106).

²¹ Ibid: 107.

²² [Blinded for peer review.]

AWS would be *needed* to address Nagel’s concern that a state waging war not “tak[e] aim [at enemy combatants] through the mundane life and survival of their countrymen, instead of aiming at the destruction of their military capacity” (1972: 139). One cannot be accused of violating Nagel’s requirement if one completely avoids civilian casualties. In a rejoinder, Sparrow contends that it is a mistake to focus solely on the non-combatants spared by AWS in deciding whether AWS count as expressing ‘respect’ for those they target. Some weapons, including AWS can be properly considered *mala in se* (evil in and of themselves) because they fail to acknowledge the humanity of those they target, and because they increase the horror of war (2016: 401). We find Sparrow’s support for these remarks unconvincing.²³ Other forms of warfare considered to be *mala in se* (e.g., starvation, poison gas, biological warfare, flamethrowers, napalm) are *cruel* to those they target. They cause gratuitous suffering, and for that reason can accurately be said to “attack the men, not the soldier” (Nagel 1972: 141). This feature is not shared by AWS. As articulated, Sparrow’s objections to AWS miss the mark. Still, we will ultimately try to vindicate part of Sparrow’s arguments. We agree with Sparrow that the inability of AWS to acknowledge the humanity of their targets poses a moral problem for their deployment, but Sparrow does not quite put his finger on the source of the problem.

Acting for the right reasons

A second important deontological objection to the use of AWS is such robots’ putative inability to act for reasons. The just war tradition demonstrates a concern with the reasons for which agents act both *ad bellum* and *in bello*. This is clear from the right intention criterion of *jus ad bellum*, but the commitment also manifests in discussions of the Doctrine of Double Effect and in the

²³ Sparrow’s derives support for his view from survey data (2016b: 402) about negative public feeling about AWS. While we do not wish to dismiss such data as irrelevant to the morality of deploying AWS, we believed it must be critically assessed before it is judged decisive evidence for the view that AWS are *mala in se*.

attention Christian just war thinkers have consistently paid to the proper orientation of the soldier’s will when committing violent acts.²⁴ This concern—with the intentions of individual actors *in bello*, not just policymakers *ad bellum*—survives through Nagel’s aforementioned “War and Massacre” up to the present day. Until AI become the full-blooded moral agents that Sparrow has in mind, it’s implausible to think that artificial intelligences could *act for reasons at all*. On the most plausible views of what it means to act for a reason, acting for a reason requires mental states such as desires and beliefs, which in turn likely require some kind of phenomenal consciousness.²⁵ Even the most sanguine of artificial intelligence researchers think that replicating this robust mental capacity in AI lies decades in the future, if it is possible at all.²⁶ (We will consider below some deontological concerns with using even highly sophisticated AWS of this sort.) If this is correct, AWS cannot act for reasons. *A fortiori* they cannot act for the right reasons when killing in war. However, notice that this objection cuts both ways: if AWS cannot act for the right reasons, then neither can they act for the *wrong* reasons. Whether this objection is successful depends on how we ought to interpret the purported requirement that soldiers act for the right reasons. Moreover, the success of this objection ultimately depends on the reasons for which most human soldiers act in deciding whom to kill. If most human soldiers act from evil motives in wartime, then a moral concern for not acting for the wrong reasons would compel us to welcome human replacement by AWS.

²⁴ See, for example, Augustine’s letter 189 to Boniface, §6. See also, [blinded for peer review]: “Augustine believes there are moral requirements for soldiers themselves to act for the right reasons. Aquinas, for his part, quotes Augustine approvingly in his *Summa Theologica* (1920). According to Reichberg, Aquinas is principally concerned with ‘the inner dispositions that should guide our conduct in war’ (2010: 264)” (2015).

²⁵ [Blinded for peer review] develop this objection. They cite Davidson (1964 and 1978) as a proponent of the desire-belief model. They cite Darwall (1983), Gibbard (1990), Quinn (1993), Korsgaard (1996), among others as proponents of the taking as a reason model.

²⁶ See Bostrom (2014) and [blinded for peer review].

2 Mutual Liability to Harm

So far we have catalogued what we take to be the major deontological objections to the development and deployment of AWS and we have identified salient weaknesses. In the remainder of the paper we develop a new deontological objection to AWS. In this section we offer an account of the moral foundation of the mutual liability to harm by combatants in wartime. In sections 3 and 4 we consider whether this moral foundation renders impermissible the use of AWS that target and kill enemy combatants. In section 5 we explore whether this objection underpins the aforementioned deontological objections, namely (a) that the use of AWS leads to responsibility gaps, (b) that it expresses a profound disrespect for enemy combatants, and (c) that AWS cannot act for the right reasons in deciding whom to kill.

Our starting point for the discussion in this section is the doctrine of the moral equality of combatants (MEC).²⁷ If we wonder whether the killing of human combatants by machines is morally permissible, a natural place to start is with an account of what makes mutual killing permissible between enemy human combatants. We will sketch a brief account of the dynamic whereby voluntary (i.e. non-conscripts) combatants can be modeled as agreeing to equal belligerency rules. (A more complex argument for why conscripts have to accept this regime will be begged here.) There is a robust debate amongst just war scholars regarding the viability of MEC as a moral theory. The view of some of the authors of this article is that the traditional, post-Westphalian construal of MEC--wherein conventional combatants on both sides of an interstate war²⁸ have equal permissions to engage in norm-guided military violence and equal liabilities to being targeted with the same--is morally correct. The arrangement is not just a convenient legal regime but reflects what some call the “deep morality” of war.

²⁷ See [omitted for blind review].

²⁸ MEC also covers privileged irregular combatants like insurgents who wear uniforms, carry their arms in the open, and obey the norms and laws of war.

We clearly do not have space here for a full-blown defense of MEC and a refutation of the doctrine’s “revisionist” critiques.²⁹ We think our argument can proceed without such a defense though, because our aim is to provide a foundation to the intuitive aversion some apparently feel to the idea of AWS killing human beings. Our suspicion is that the intuitive aversion to killing by AWS is rooted in how service personnel—and perhaps the general public—understand the profession of soldiering.³⁰ If the aversion is linked to MEC, people who subscribe to MEC will feel this aversion regardless of whether MEC is morally correct. The account of MEC presented below will therefore not be presented with a hope of convincing skeptics, but of providing precise language for articulating the intuitive deontological objection. We understand that ours will be a controversial account but it seems to ably explain why many find AWS grotesque.

An Account of the Moral Equality of Combatants

A joint right is a moral right that can only be enjoyed in a collective setting, like a right to collective security or a right to political determination. Because joint rights can only be satisfied in a collective setting, joint rights engender collective responsibilities. People partially fulfill their collective moral responsibilities to meet joint rights by creating just institutions, including governments, that protect those rights. Once these institutions are established, collective moral responsibility is largely transferred to institutional actors whose properly-constituted institutional imperatives become moral norms for those actors because of their institutional role in protecting joint rights.³¹ For example, police hewing to their properly-constituted professional imperatives are fulfilling moral duties, not merely “doing their job.” People outside the institutions

²⁹ One of the authors advances that full-blown defense and critique in [blinded for review].

³⁰ One of the authors has spent a career working with the US military. In his experience, service personnel overwhelmingly accept MEC for conventional adversaries.

³¹ Seumas Miller, (2010: 57, 68, 77, 80).

must support the (just) institutions when those institutions are the most effective instruments available to protect others' joint and individual rights.³² One meets this duty to support just institutions by paying for their maintenance, complying with the rules of the institutions, and refraining from attempts to subvert them. The duty also entails ceding certain claim-rights to the institutions' professionals (explained below).³³ As with other duties, compliance respects others' rights but also contributes to a reciprocal system ultimately protecting the rights of the duty-bearer. So, for example, paying tax to support one's local police department ideally helps protect one's neighbors' rights as well as one's own.

One owes a duty to the inhabitants of a state to uphold just institutions and "pays" what one owes to the government employees who are those inhabitants' agents. A duty to deliver X to Y means one cedes to Y a claim-right against certain forms of treatment that are necessary for the provision of X. In just states, a duty to uphold just institutions means ceding certain claim-rights against otherwise rights-violating activities of state agents, where 1) those state agents are competently pursuing their obligations and duties, and 2) when insisting on one's rights would otherwise prevent state agents from serving their principals. So for example, one indirectly meets one's duty to respect and protect others' rights by supporting law enforcement institutions and ceding certain claim-rights to the (norm-compliant) police who protect one's neighbors. This ceding

³² Miller (2010: ch. 2), Rawls (1999: 351–4), and Waldron (1993: 3–30, 26, 28) derive the duty differently. (Cf. Nozick 1974: 102, 110). Support ought to be withdrawn if the institutions become significantly corrupt and fail to meet the goals for which they were established over a long period of time. Yet the anarchic risks of non-compliance with institutional rules are significant so support for institutions generally oriented to morally valuable goods is indicated even in the event that the institution pursues a particular unjust project. For example, tax-payers should not cease paying tax to support public schools because of a dubious curriculum implemented one year. Citizens should not overthrow their governments because of a bad foreign policy decision.

³³ Waldron, 8-10.

of certain claim-rights gives the police liberty-rights in turn, creating the space for them to perform their norm-compliant actions without wronging the affected parties. This dynamic explains how police in a just society can arrest, detain, and non-violently interrogate someone in their jurisdiction without wronging him, so long as they hew to the norms of due process. Incompetent, brutal, or corrupt police activities are not covered by this permission. Without this ceding of rights, police would be morally prohibited from investigating and arresting suspicious persons, and the rights of potential and actual crime victims would go unprotected. In this way, police benefit the very inhabitants who can be modeled as ceding rights to them.

Importantly, the duty to support just institutions is not restricted to institutions of one's own state; it extends in different ways to foreign institutions. The duty to support just institutions is derived from the more basic duty to aid and protect the rights of other human beings—a duty all people (capable of being duty-bearers) owe to all other people, regardless of national identity or citizenship. So, for example, one cedes claim- and liberty-rights to the state agents of a foreign state when insisting on those rights would morally impede foreign agents from performing their duties to protect their own citizens, residents, and guests. For this reason, a tourist's rights are not violated when she is compelled to obey the commands of foreign police officers when abroad, when those officers are acting as agents of just institutions and conducting themselves in just, norm-compliant ways.

Normally, the duty to support just institutions imposes few demands on a person in her home state for the benefit of people in other states. In war-time though, the duty to support just foreign institutions can serve as the ground for the moral equality of combatants and egalitarian military norms. In war, the duty to support just foreign institutions entails ceding claim-rights against foreign agents acting according to their professional duty. Combatants then are not wronged when targeted by norm-compliant enemy combatants representing just political entities. Professionals are duty-bound to adhere to professional norms guiding professionals to achieve the morally valuable goals of their

professions in a way optimizing trade-offs between the efficient pursuit of those goals and rights infringements to those affected in the pursuit of those goals. Service members can be modeled as ceding claim-rights against being targeted with lethal violence by enemy combatants according to the terms of military norms optimizing a balance between maximizing the one military's interests while minimizing suffering to their enemy. This is beneficial to all military personnel insofar as it enables them to fulfill their duties by achieving the morally valuable goals of their professions, *qua* agents responsible for upholding just institutions. For example, the principle of necessity benefits service personnel *qua* agents of just institutions by permitting them to target enemy personnel whom they need to target in order to seize an objective. The principle of necessity also benefits service personnel *qua* patients by sparing them from gratuitous violence. Combatants can be modeled as ceding these rights to norm-compliant adversarial personnel because of their duty to support just institutions, but also because they benefit from the adversarial belligerent practice of which these actions are a part. The members of both militaries are benefited by the adversarial practice of which they take a direct part because the practice permits their own state to secure its legitimate national security goals without risking undue suffering on the part of its noncombatants and combatants.

If service members did not cede claim-rights against being directly targeted by norm-compliant enemy personnel representing basically just states,³⁴ foreign military personnel would be wrong to serve their state in the same way the first military was permitted to serve its basically just state. *Ex ante*, there is no basis on which to model one group of service members' duty to protect their political entity while withholding permission to another group performing the same duty. We cannot model combatants' consent to a regime which only extends belligerent privileges to the just side since there is no international body with enforcement mechanisms which can definitively say which war is just. States are in a quasi-anarchic international system in which political leaders have to (ideally)

³⁴ Basically just states largely respect the basic rights of their inhabitants and equitably enforce the law; they need not be democratic.

make their best judgment about when war is warranted to protect their state's interests. An asymmetrical regime only permitting combatants on the just side to fight would be unworkable as there would be no clear authority to consult about the justice of a war.

It is important to stress that service members' duty to observe their professional norms and their liability to being the direct target of norm-compliant enemies is not based on their express consent. On the picture we are describing, by voluntarily enlisting, enlistees accept the role of service members and take on the norms attendant to that role in an adversarial system. Service members cannot be modeled as waiving their rights to life per that role—as this would permit anyone to kill them. They do not expressly waive their claim-rights against being killed,³⁵ but have to be modeled as ceding claim-rights against being directly *targeted* by *combatants* representing basically just states with lethal action conforming to military norms. Service personnel in military A are not left defenseless by ceding these rights because their enemies in military B also cede claim-rights against being targeted by personnel in military A. Service personnel therefore gain reciprocal liberty-rights to directly target their enemies. This permission creates the moral space for them to perform their duties in service of their political entities without wronging their combatant targets. Importantly, combatants can still be wronged by being wantonly killed, killed in ways that impose unnecessary levels of suffering, or killed by unprivileged irregular militants and ordinary noncombatants (so long as the combatants confronted by the noncombatants are themselves hewing to their military norms). So, MEC still allows for cases in which combatants are wronged by being killed in combat. In the remainder of this essay we consider whether MEC entails that combatants would also be wronged if killed by an AWS.

While this essay is focused on combatants, some might also wonder if AWS can wrong noncombatants by killing them. To that end, let us briefly discuss noncombatants' rights. Noncombatants also have a duty to support just foreign military institutions because

³⁵ Cf. Hurka, 210.

of their duty to respect the rights of foreigners, including those rights best protected by militaries. This means noncombatants in country A cede certain claim-rights that would otherwise morally prohibit foreign service personnel in country B from protecting their people to the same extent service personnel_A protect noncombatants_A. Since personnel can rarely fight a war without incurring noncombatant casualties, all noncombatants can be modeled as demanding their personnel fight just wars in order to protect their rights even while risking foreign noncombatant casualties. The alternative would be contingent pacifism. Yet since no one wants to be collaterally killed in a necessary military attack, all can be modeled as accepting for example, the military norm of proportionality, which permits the foreseen but unintentional killing of noncombatants but only when those casualties are the unavoidable cost of securing a tactical victory the impartial goodness of which outweighs the badness of the collateral harm. Proportionality maximizes the goods that all noncombatants want and minimizes the harms that none want. This means that noncombatants who have a military fighting on their behalf are not wronged if they are collaterally killed by enemy personnel scrupulously adhering to the norm of proportionality.

3 Ceding claim-rights against being targeted to AWS?

In this section, we investigate what ceding a right involves in order to determine if service personnel can be modeled as ceding to AWS their claim-rights against being targeted with lethal violence. We propose that they cannot. To claim otherwise is to commit a kind of category mistake. A member of class X cedes claim-right R to members of class Y, giving members of Y a liberty-right to act in ways that would otherwise violate R. That a member of class Y has a liberty-right, a permission, to materially infringe R (that again, would otherwise violate R), means that a member of Y may permissibly choose to materially infringe R.³⁶ Because

³⁶ To violate a right is to wrongfully infringe that right.

permissions only apply to duty-bearers, this permission presupposes that the permission-holder is capable of being a duty-bearer.

Having the capacity for moral responsibility is a necessary, but not sufficient, component of being a duty-bearer. An agent with the capacity for moral responsibility (1) is able to voluntarily engage in intentional action; (2) is physically free to act or refrain from acting; (3) has a basic understanding about how her actions affect people and things; (4) can form an intention regarding a specific action and carry it out; and (5) can form her intentions in response to her recognition of reasons for action (be they moral or non-moral reasons). Those who are very young, mentally-ill, under hypnosis, or under the influence of mind-altering chemicals typically lack some of these components of responsibility.

In addition to having the capacity for morally responsibility, a duty-bearer must be *morally mature*. Morally maturity includes (6) the ability to largely control one's impulses; (7) having morally significant interests; (8) having an appreciation of the rights and duties relevant to her and others' interests; (relatedly) (9) the ability to judge when she must restrain the pursuit of her own interests in accordance with her duties; and (10) the ability to weigh whether it is appropriate to exercise her own rights and privileges when she is not duty-bound to omit actions but when exercising her rights will affect others.³⁷ Moral maturity is necessary to stand in a reciprocal relationship with duty-bearers—to be trusted with liberty-rights gained from relationship partners' ceded claim-rights. For example, regarding (9), a landlord understands she is not permitted to build an addition on a building she does not own, but rather, has a duty to defer to the second building owner's rights. She can understand and endorse that duty because she expects the other landlord to observe the same duty toward her property. Regarding (10), the

³⁷ As robots do not experience impulses of the relevant sort, we will ignore condition (6) for the rest of this essay. We also cannot answer whether an AWS can choose to omit the pursuit of its own rights in order to defer to others' interests without answering the key question of this essay. That question is whether AWS have liberty-rights to kill enemy service members. We will therefore put element (10) aside as well.

landlord may choose to delay evicting a tenant in her *own* building—in accordance with her rights—if that tenant’s delinquent rent is due to the onset of a serious medical condition.

The reader will detect Aristotelian and Kantian echoes in condition (5), that duty-bearers are capable of recognizing and responding to moral reasons: moral maturity requires more than the avoidance of material rights violations. A comatose person violates no one’s rights, but is not a duty-bearer. A little boy who refrains from hitting his brother might be praised for obedience rather than moral maturity because he omits violence merely from fear of consequences. Moral maturity requires compliance with outward expectations for self-elected moral reasons. Understanding and being able to weigh moral reasons necessitates appreciating the value of human life and well-being, autonomy, and associated moral concepts like rights and duties. If one fails to possess a concept of these considerations, then one cannot understand their significance or make judgments about the reasons that derive from them.

Deontological explanations of the conditions under which rights can be ceded presuppose reciprocity amongst rights-holders and duty-bearers. One cedes rights to moral *equals*, agents with the same capacity for duty-bearing.³⁸ An adult does not cede rights to

³⁸ In order to raise questions about AWS, Sparrow invokes Thomas Nagel’s insistence that the moral basis of military violence has to be an interpersonal relationship between subjects, (Sparrow 2016a). Since an AWS is incapable of an interpersonal relationship, it cannot engage in permissible killing. Sparrow may be missing the force of Nagel’s argument, which is focused on the recipient of military violence rather than the agent (1972). The recipient of violence is treated with respect when he is targeted for something he chose to do, like becoming a combatant, as opposed to something that has nothing to do with his subjectivity, like his ethnic affiliation or his presence in a certain area. The use of indiscriminate weapons is disrespectful because such weapons do not distinguish between people based on their status or activities. Writing in 1972, Nagel certainly must be assuming that a human agent is engaged in discriminate targeting, but it seems that a sophisticated AWS could engage in that kind of distinction, only targeting armed personnel or military materiel instead of targeting all people in an area.

children, adults with certain kinds of serious mental illness or senility, non-human animals, or machines.³⁹ None are duty-bearers.

We might end the essay here by arguing that human combatants cannot be modeled as ceding claim-rights to robots against being targeted with lethal violence. One does not, and cannot, cede rights to a toaster, drill, power lathe, or a plastic injection molder. Watches, phones, and computers enjoy no claims over humans. Such conventional machines lack agency, the capacity to know, form intentions, have serious interests, restrain themselves, or understand rules *as* moral rules. They are mere tools for their human agents. Therefore, since it is a machine, no AWS can have permission to kill service personnel or noncombatants. From this observation, one can show that it would also be impermissible for a human service member to deploy AWS against human targets. A duty-bearer's permission to harm enemy combatants does not include outsourcing the task of harming to an entity to which the enemy has not ceded and cannot cede his right against being harmed. Let us consider some civilian comparisons involving lesser material infringements. A client cedes privacy rights against her tax accountant's viewing her financial documents, but the accountant's permission to view confidential documents does not include a permission to send those documents to another firm for assessment. A business executive engaged in negotiations cannot transfer his decision-making authority to his seven-year-old daughter.

This account of the moral objection to the use of AWS allows for a principled distinction between the use of AWS and the use of non-autonomous lethal weapons. One service member cedes a claim-right to another to target him with lethal non-autonomous weapons. This is because non-autonomous weapons do not target enemy combatants. When a soldier fires a gun at an enemy combatant the soldier, not the gun, targets the combatant. The service member is not subcontracting her agency to a non-autonomous weapon even if it travels a long distance from the agent.

³⁹ One can also enjoy rights by virtue of natural properties, not strictly through having rights ceded to one. So babies might have rights despite not being eligible to be moral contract partners.

Yet ending the essay here would be unsatisfying for two reasons. First, advocates of AWS might argue that we are begging a question against them if we are relying on an argument equating future AWS with toasters. Perhaps future AWS will be so close to humans in relevant ways that we could think of them as morally responsible agents capable of bearing duties. It is hard to be definitive here because there is a divergence among thinkers regarding the conditions necessary for moral responsibility and the requirements for duty-bearing as well as uncertainty about what the future holds for AWS. Second, we are trying to provide the basis of a powerful but inchoate deontological objection to AWS killings; asserting that something lacks permission to do something seems to under-explain the intuitive horror or offense of being hunted and killed by a robot. There seems to be something else going on besides the mere impermissibility of the killing, which would also apply to being killed by an unprivileged belligerent like a terrorist. Yet the prospect of being killed by an unprivileged belligerent, even one who fights dirty by posing as a noncombatant, seems to conjure a very different emotion than being killed by a robot.⁴⁰

4 Can humans cede rights to *sophisticated* AWS?

Could a service member be modeled as ceding his rights not to be targeted with lethal violence to a so-called sophisticated AWS? Let us assume a sophisticated AWS can discriminate between available targets; discern visible cues indicating a person's combatant status, including signs of surrender; and adhere to *jus in bello* and the law of armed conflict when determining targets. We will be charitable to roboticist optimists and grant that a future, sophisticated AWS could have features that are at least *functional* equivalents to the components of human moral responsibility. That is, we grant that AWS would act just like a morally responsible agent would act.

⁴⁰ Full disclosure, one author of this paper does not share this asymmetrical aversion to death by robot compared with death by terrorist.

Yet it is hard to imagine how a service member can cede his rights not to be targeted with lethal violence to a so-called sophisticated AWS, because it is hard to imagine how an AWS that is (merely) functionally equivalent to a morally responsible agent can be a bearer of duties. Recall the relevant components of being a duty-bearer: having important interests, having the capacity to judge whether countervailing rights generate reasons against an action, and having the capacity to judge whether it is appropriate to act within one's rights by frustrating others' interests.⁴¹ An AWS can, in a sense, restrain its own projects (e.g. its programmed "goal" of targeting the enemy) in deference to noncombatants' rights. An AWS can be programmed with moral rules, but it is difficult to conceive of how it can obey the rules while understanding them *as* moral rules. An AWS cannot do what any mentally normal adult can do, which is distinguish moral rules in kind from rules of a game or rules of small boat navigation or rules for matching shirts and ties.⁴² It would seem that an AWS cannot appreciate a set of rules as moral rules because an AWS cannot appreciate the value of human life or the authority of moral rules. Intuitively, appreciating the authority of moral rules requires understanding the rules as generating *reasons* for action. An AWS cannot do this, and it certainly cannot inform its targeting decisions on the basis of judgments about the reasons generated by those rules. To invoke the language used in section 3, AWS are morally immature and, hence, are not candidate duty-bearers.

Let us consider in more detail the claim that a sophisticated AWS cannot understand the value of human life. There are two key aspects of human life a machine will never experience, which allow

⁴¹ Some utilitarians would argue that agents have duties despite the non-existence of rights. Since most contemporary just war theorists operate in a deontological idiom, we will confine our discussion to that broad moral framework.

⁴² Psychopaths, famously, have trouble either distinguishing moral rules from non-moral ones (Borg and Sinnott-Armstrong, 2013), or else feeling the force of the moral rules they do recognize (Cima et al, 2010). See these, among other entries in the debate over the moral psychology of psychopathy.

any mentally-normal adult to appreciate the value of human life.⁴³ A mentally normal adult can appreciate his own mortality and the value of his own experiences and life projects. Understanding the moral gravity of human life requires, as a starting point, an appreciation of how much another person's life (and the goods that it contains) matters to *him*. To use Robert Nozick's phrase, it requires appreciating the significance to him of the fact that "His is the only life he has."⁴⁴ But this appreciation is possible only if I first experience the significance of my own life to *me*. Typically, this valuation is connected with an appreciation of mortality. To this point, it is often commented that teenagers are apt to behave recklessly because they think they are invincible. Part of maturity is a subtle appreciation of one's own limits; a humbling that comes from a diminishment of physical powers; and perhaps, the imaginative and empathetic capacities needed to appreciate the position of the dead and dying. As one gets older, one comes to not merely cognitively grasp—as children can—but *appreciate, affectively*, that, just as grandma and grandpa, and Bob's grandpa, and that poor kid in 12th grade who died in car accident, that *I will die, too*.

Grasping the value of human life also requires that I can *value* and *feel* the weight of my projects and commitments. This valuation of projects comes from the experience of choosing these projects over others; the effort one expends to pursue them; and the pain one feels over failed endeavors. The reasons one values one's commitments to projects, causes, people, and ideas is very complex, but in part comes from their centrality to one's identity. They are important because they make one who one is. One's identity is made up of an intricate web of commitments, experiences, knowledge, emotions, and aspirations. A moment's deliberation reveals that *everyone else* is just as complex—good and bad, brave and weak—

⁴³ Discussions of the ethics of AWS sometimes invoke vexed terms like "the value of human life," "moral weight," "moral gravity," and so on. We want to be careful about the use of such evocative but imprecise terms lest arguments reduce to "I know it when I see it"-style appeals that beg questions against proponents of AWS.

⁴⁴ Nozick (1974: 33).

as dense and unique a creation as any pointillist work of art. So, it is fairly easy to conclude, *their lives are just as valuable as one's own*.

It seems incredible to us that an AWS could have this capacity to subjectively value its own existence and affectively project this value on the part of other people (or sophisticated robots), partly because we are skeptical that an AWS could feel anything at all. An AWS can perhaps be programmed to recognize the difference between a living and a dead person; it can be programmed to obey *jus in bello*; it can pass up potential targets that do not comply with the targeting criteria. An AWS can perhaps even surpass its initial programming by adapting to enemy tactics. Yet we cannot conceive of how an AWS can have a sense of its own history and its mortality. Even if it logs all its actions, it cannot cast its eyes over them in order to have a sense of self. It cannot fear death nor savor life. It cannot affectively project these qualities on the part of its human targets. It therefore cannot take life with an appreciation of the gravity of what it is doing. Again, even a sophisticated AWS would be morally immature and hence not a duty-bearer. In this way, an AWS is in a worse position to appreciate the moral gravity of human life than the sociopath. For at least the sociopath appreciates the value of his life to him. He is simply unable to extend this concern to the lives of others.

In a deontological idiom, people are modeled as ceding rights to, or demanding respect from, moral equals. Whether two individuals are moral equals is determined by whether they are endowed with the same basic moral capacities. The relevant moral background knowledge necessary to honor the moral power granted to the permission-holder by the rights-ceder includes an appreciation of the value of human life—the value of one's own life and related moral interests, and the felt recognition of that symmetrical value in others.⁴⁵ A service member can be modeled as

⁴⁵ Peter Asaro argues that the application of morally rich laws cannot be automated because such laws are designed to be interpreted by people. For example, the right to due process is essentially a right to "question the rules and the appropriateness of their application in a given circumstance, and to make an appeal to informed human

ceding a claim-right against being targeted with lethal violence to another combatant who understands the gravity of what he is doing, but not to an entity incapable of understanding this gravity.⁴⁶ A wounded (volunteer) service member cannot rationally resent his conventional enemy assailant because the wounded service member freely entered into an adversarial, norm-governed, reciprocal relationship with his state's future enemy combatants wherein he can be modeled as granting them liberty-rights to try to kill or wound him. Yet he cannot be modeled as ceding any rights to a *robot* and so the wounded service member *can* rationally resent the enemy officer who deployed the AWS against him. The enemy officer breaches the reciprocal relationship that gives both sets of service personnel in a war belligerent rights and liabilities. This, we believe, is the intuitive foundation of Sparrow's view that using AWS expresses a profound disrespect for the enemy. AWS-deploying officers do not see their enemies as moral equals worthy of a reciprocal relationship.

Noncombatants can be modeled as ceding claim-rights against having their lives endangered in the course of other agents pursuing their goals as a way of supporting just foreign institutions and as a reciprocal cost of demanding that their service personnel engage in activities endangering foreign noncombatants. Similar to

rationality and understanding" (2012: 700).

⁴⁶ This argument also addresses Michael Robillard's argument that the deontological concerns about AWS are misplaced because an AWS does not make genuine decisions of its own but merely acts on conditional orders programmed into it ahead of time by human beings. On Robillard's view, the designers of the AWS are therefore responsible for the deaths directly incurred by the AWS (forthcoming: 6-7). A complicated technical and philosophical discussion would be required to address whether a given level of complex machine learning could result in an entity able to make genuinely autonomous decisions distinct from one of the conditional prompts programmed by the AWS engineers. To our point though, combatants do not enter into reciprocal moral relationships with weapon designers but rather with the agents who decide to use those weapon systems. If weapon designers did cede their claim-rights against being targeted by the combatants their weapon systems threaten, then, counter-intuitively, an elderly, retired engineer could be permissibly targeted in war-time by combatants threatened by aircraft the engineer contributed to the design of 30 years prior.

the above argument, this ceding of claim-rights can only be to duty-holders who understand the value of human life and so will engage in proportionality calculations with an appropriate reluctance. Again, the idea is that moral respect requires more than deploying a non-duty-bearing robot that can perfectly mimic the morally appropriate actions. On our view, an AWS and human who determined the same number of collateral casualties to be permissible, and then engaged in a targeting maneuver on the basis of this calculation, would not be performing the same action.

Similarly, service personnel or noncombatants would be wronged if the enemy deployed an army of trained dogs, or of children, or of mental patients suffering from violent hallucinations, or of sociopathic killers who derive sadistic pleasure from killing. A service member or noncombatant cannot be modeled as ceding rights to members of these groups because they are not duty-bearers. They are not duty-bearers because they lack one of the capacities necessary for qualifying as a duty-bearer: either they cannot see their enemy as *morally* valuable; they engage in no moral deliberation regarding their moral reasons; or they feel no moral tension over the action.⁴⁷ By contrast, the duty-bearer has the capacity to suffer over the decision to take a life.⁴⁸ She might need to take a quiet moment to herself after the act, or talk about it later with sobriety and agent-regret.⁴⁹ The duty-bearer need not know the target's biography in order to feel this weight but needs to know that he is a human being with the same basic moral value as the agent.

One concern with long-range killing is that it is easier to avoid experiencing the moral value of human life because the agent does not see the effects of his action. Yet a missile officer on a submarine nonetheless makes the targeting decision that he has been

⁴⁷ This argument is consistent with Nagel (1974: 136).

⁴⁸ Granted, an experienced service member may be able to kill later in his or her career without much emotion.

⁴⁹ “Agent regret” is a term introduced by Bernard Williams (1993), referring to the emotion one feels following one’s non-culpable causation of harm.

granted a right to do, and he is a duty-bearer. Because he is a duty-bearer, he can still intellectually grapple with the weight of his actions, despite the distance between him and his target. An AWS of the sort we are considering here categorically cannot do so. Part of what makes the action weighty for the missile officer is the understanding that he and his fellow citizens are in principle vulnerable to an enemy's attack as well, and that the foreign combatants' and noncombatants' lives areas significant to them as the officer's life is to him.

To conclude this section, some have suggested that the deontological concerns with AWS such as the epistemic uncertainty of the authorizing human agent, the metaphysical indeterminacy of the AWS's victims, and the lack of a personal relationship between human agent and victim, also speak against widely used long-distance, human controlled weapons. The widespread acceptability of such human-controlled weapons undercuts the idea that there is a genuine and novel moral concern with AWS killing. The relevant moral difference between AWS and long-distance weapon systems is that AWS targets combatants for harm on its own. Yet, as we have argued, AWS cannot be the recipient of liberty-rights gained by humans ceding claim-rights to them. AWS lack a moral sense necessary to form reciprocal relations between combatants. We can conclude that AWS lack the sort of permission to kill service members in war that human service personnel can gain.

It will not do to say that the human agents who deploy AWS are the recipients of the liberty-rights, because the intermediate status of AWS agency renders it impossible for those human agents to count as targeting combatants targeted by the AWS they deploy. A human service member's permission to harm does not include a permission to deploy AWS against human targets, because a duty-holder's permission to harm enemy combatants does not include a permission to outsource the task of harming to an entity to which the enemy cannot cede his right against being harmed. It is for this reason that the victims of AWS attacks are disrespected by these attacks.

This last point requires some further elaboration, for someone might object to the foregoing discussion by presenting a dilemma: either AWS are sophisticated enough to adhere to the moral rules of armed conflict or they are not. If they are not sophisticated enough to do this, then AWS clearly aren't agents, but they *also* should not be understood as killing the enemy. In this case, it's simply unclear why we should think of using an AWS to kill people as "outsourcing the task of harming to an entity" at all. It's the enemy combatant that kills people and the weapon is merely the means by which they do so.⁵⁰ However, if AWS *are* sophisticated enough to be able to adhere to the moral rules governing armed conflict, then they must possess attributes sufficient for moral agency and hence duty-bearing. For example, one might think that in order to make proportionality calculations, AWS *must* be able to appreciate the moral value of human life and the moral reasons this value generates. If AWS are moral agents, then it seems that enemy combatants *could cede to the AWS* their right against being targeted. Whichever horn of the dilemma we choose, the objection goes, the MEC does not pose a moral problem for the deployment of AWS.

We grant the point about unsophisticated AWS. In reply, let us focus on the future, highly sophisticated AWS, which can adhere to the moral rules of armed conflict at least as well as human combatants. Would the capacities possessed by such a machine be of the sort that would make it qualify as a moral agent to whom combatants can cede their claim rights against lethal targeting? We think not. The ability of an individual to adhere to the principles of just war theory and the laws of armed conflict depends on the interaction of a number of capacities, only some of which are relevant to moral agency. Some of the current authors [citation blinded for peer review] identify three types of mistakes that cause combatants to violate the principles of just war: empirical mistakes; practical mistakes; and genuine moral mistakes. Empirical mistakes are mistakes in discovering and identifying what the empirical facts are. To misidentify a potential target as carrying a gun when the

⁵⁰ We thank an anonymous referee for the *Journal of Applied Philosophy* for pressing us on this point.

object is in fact a camera is to commit an empirical mistake. Genuine moral mistakes, are mistakes in moral judgment, how to weigh the relevant moral considerations once they have been discovered, etc. These mistakes occur when we come to the wrong normative answer about a moral problem, even given full information about the empirical facts. Finally, practical mistakes can occur even when an individual has correctly identified the relevant empirical and moral facts, but fails in responding to those facts (e.g., by reacting a moment too slowly, or by missing one's target and shooting an innocent person due to mental fatigue). This tripartite distinction between empirical, moral, and practical mistakes can help us to see how it is possible that a sophisticated AWS could be at least as good as human combatants at adhering to the moral rules governing armed conflict while nonetheless lacking the capacities sufficient for moral agency. Consider our previous discussion of this point:

There is good reason for thinking that AWS could commit drastically fewer empirical and practical mistakes than human soldiers. Decisions on the battlefield must incorporate massive amounts of data, and they must be made in seconds. Adams (2001) points out that the tempo of warfare has increased dramatically in recent years, and so it will presumably only accelerate further. The human mind is only capable of incorporating so much information and acting on it so quickly. The day may come when human combatants simply cannot respond quickly enough to operate effectively in modern warfare. AWS has the potential to incorporate massive amounts of information, thereby avoiding the empirical and practical mistakes that humans are eventually bound to make in the ever-quickening pace of battle [citation blinded].

A sophisticated AWS might adhere at least as well as human combatants to the principles of just war because it makes fewer empirical and practical mistakes, even if it lacks some of the capacities necessary for moral agency. Such a machine, we propose,

could target enemy combatants in a way that counts as “outsourcing the task of harming” *without* possessing the capacities necessary for moral agency. Therefore, the objection by dilemma fails against our argument. As we wrote in the Introduction, our aim in this paper is not to conclusively argue against the use of AWS in all cases. Our argument does provide, to our eyes, a genuine and novel deontological objection to AWS killing.

5 Vindicating moral objections from intentions, responsibility-gaps, and the interpersonal relationship between combatants

Earlier we claimed that the novel deontological objection to AWS proposed in this paper could be shown to underpin other deontological objections to AWS that had been raised in the literature. In closing we hope to make good on that promise. Though each of the objections we surveyed in section 3 misses the mark, each objection is, in some way or another, vindicated by the novel objection we have identified—that, according to the MEC, it is impermissible to deploy AWS to harm enemy combatants, because enemy combatants cannot cede rights against being harmed to AWS.

Sparrow objects to the use of AWS because the use of AWS would leave us with no one to appropriately hold accountable for the enemy combatant deaths it causes. One worry for this objection was that it is unclear why it matters that there be someone to hold accountable for the deaths of enemy combatants, if AWS adheres perfectly to the *jus in bello* principles of discrimination and proportionality. Our objection, based on MEC, can fill the explanatory gap: if there is no one to hold accountable for the deaths of enemy combatants, then the agent of their death is not a duty-bearer to whom they have ceded their claim right against being targeted. Being killed is thus a violation of their rights.

Sparrow also objects to the use of AWS because it is a weapon *mala in se* in virtue of failing to acknowledge the humanity of combatants. Employing a weapon that kills without acknowledging the humanity of its victims is disrespectful of those

victims. The problem for this objection is that there is another clear sense in which, if AWS will be better than human personnel at discriminating between legitimate and illegitimate targets, employing AWS in combat is *highly respectful* of its victims: it provides greater assurance to targets of aggression that, should they be rendered illegitimate targets because of surrender or incapacitation, they will not be unjustly killed. Here too our objection, based on MEC, can close a hole in Sparrow's argument. The sort of disrespect that is expressed when one employs AWS is not ameliorated by observing that AWS will kill all and only legitimate targets, because AWS are not authorized to kill in the first place. AWS are not duty-bearers, and so combatants cannot cede a claim-right to AWS against being targeted by them. All AWS targets are illegitimate targets. This is the fundamental reason that AWS's inability to acknowledge the humanity of its victims renders its use disrespectful of those victims.

Finally, two of the authors have elsewhere argued that using AWS is morally objectionable because AWS are unable to make moral judgments or to act for moral reasons in deciding whom to kill. The trouble for this objection was that it cuts both ways: if AWS cannot act for the right reasons, then neither can they act for the *wrong* reasons. So, whether the objection is successful depends on how we ought to interpret the *in bello* requirement, if there is one, that soldiers act for the right reasons. Moreover, whether this objection identifies a genuine moral difference between AWS and other widely accepted methods of waging war ultimately depends on the reasons for which most human soldiers act in deciding whom to kill. If most human soldiers act from evil motives in wartime, then a moral concern for acting for the right reasons would compel us to welcome human replacement by AWS. Our objection, based on MEC, can explain why the inability of AWS to act for moral reasons renders its use morally problematic without (a) a commitment to the view that waging war justly requires that soldiers act for the right reasons in deciding whom to kill or (b) assuming that most human soldiers act from morally good motives. On the argument propounded in this paper, it is the *capacity* to act for moral reasons that matters, not the exercise of that capacity.

The capacity to act for moral reasons is morally significant, because it is necessary to be a duty-bearer, and a combatant cedes his right to be killed only to duty-bearing enemy combatants. Because AWS cannot act for reasons, it cannot be a duty-bearer. When AWS kills a combatant, the agent of the combatant's death is not a duty-bearer to whom they have ceded their claim right against being targeted. Being killed by AWS is thus a violation of their rights.

Bibliography

- Asaro, Peter. 2012. "On Banning Autonomous Weapon Systems: Human Rights, Automation, and the Dehumanization of the Lethal Decision-making," *International Review of the Red Cross* 94.886, 687-709.
- Borg, Jana Schaich, and Walter Sinnott-Armstrong. "Do psychopaths make moral judgments?." *Handbook on psychopathy and law* (2013): 107-128.
- Bostrom, N. (2014). *Superintelligence: Paths, Dangers, Strategies*. OUP
- Burri, Susanne. (2017). "What is the Moral Problem with Killer Robots?" in *Who Should Die?*, eds Bradley J. Strawser, Ryan Jenkins, and Michael Robillard (Oxford: Oxford University Press, 2017), 163-185.
- Cima, Maaïke, Franca Tonnaer, and Marc D. Hauser. "Psychopaths know right from wrong but don't care." *Social cognitive and affective neuroscience* 5.1 (2010): 59-67.
- Darwall S. (1983) *Impartial reason*. Cornell, New York
- Davidson D. (1964). Actions, reasons, and causes. *Journal of Philosophy* 60(23):685–700

- Davidson D. (1978). Intending. In: Yirmiahu (ed) *Philosophy and history of action*. Springer, p 41–60.
- Docherty, Bonnie. (2012). “Losing Humanity: The Case Against Killer Robots,” *Report for the Human Rights Watch*: pp. 39–41.
- Gibbard A (1990) *Wise choices, apt feelings*. Clarendon, Oxford
- Guarini M, Bello P. “Robotic warfare: some challenges in moving from noncivilian to civilian theaters. In: Lin P, Abney K, Bekey GA (eds) *Robot ethics: the ethical and social implications of robotics*. MIT Press, Cambridge, 2012), 129–144
- Hurka, Thomas. 2007. Hurka, Thomas. "Liability and just cause." *Ethics & International Affairs* 21.2: 199-218.
- Kahn, Leonard. 2017. "Military robots and the likelihood of armed conflict" in *Robot Ethics 2.0* (eds. Lin, Jenkins, and Abney). Oxford University Press.
- Korsgaard C (1996) *The sources of normativity*. Cambridge University Press, Cambridge
- Krishnan, Armin. *Killer Robots: Legality and Ethicality of Autonomous Weapons* (Surrey, UK: Ashgate Publishing Limited, 2009);
- Matthias, Andreas. "The responsibility gap: Ascribing responsibility for the actions of learning automata." *Ethics and information technology* 6.3 (2004): 175-183.
- McNaughton, David, and Piers Rawling. "On defending deontology." *Ratio* 11.1 (1998): 37-54.
- Miller, Seamus. *The Moral Foundations of Social Institutions*, (Cambridge: Cambridge University Press, 2010).

- Nagel, Thomas. "War and massacre." *Philosophy & Public Affairs* (1972): 123-144.
- Nozick, Robert. 1974. *Anarchy, State, and Utopia*. New York: Basic Books.
- Quinn W (1993) *Morality and action*. Cambridge University Press, Cambridge
- Rawls, John. 1999. *A Theory of Justice*, 2nd ed., Cambridge, MA: Belknap Press.
- Robillard, Michael. Forthcoming. "No Such Thing as Killer Robots," *Journal of Applied Philosophy*. DOI: 10.1111/japp.12274, 1-13.
- Roff, Heather M. (2013). "Killing in War: Responsibility, Liability, and Lethal Autonomous Robots," in Fritz Allhoff, Nicholas G. Evans, and Adam Henschke, eds., *Routledge Handbook of Ethics and War: Just War Theory in the Twenty-First Century* (Milton Park, Oxon: Routledge,)
- Scheffler, Samuel. (1984). *The Rejection of Consequentialism: A Philosophical Investigation of the Considerations Underlying Rival Moral Conceptions*. Oxford University Press.
- Schmitt, Michael N., and Jeffrey S. Thurnher. "Out of the loop: autonomous weapon systems and the law of armed conflict." *Harv. Nat'l Sec. J.* 4 (2012): 231.
- Sharkey, Noel. "Automated Killers and the Computing Profession," *Computer*, vol. 40, issue 11 (2007), p. 122.
- Sparrow, Robert (2007). Killer robots. *Journal of Applied Philosophy* 24 (1):62-77.

- Sparrow, Robert (2015). Twenty seconds to comply: Autonomous weapon systems and the recognition of surrender. *International Law Studies* 91:699-728.
- Sparrow, Robert. (2016a). Robots and Respect: Assessing the Case Against Autonomous Weapons Systems. *Ethics and International Affairs*. 30(1): 93—116.
- Sparrow, Robert. (2016b). Robots as “Evil Means”? A Rejoinder to Jenkins and Purves. *Ethics and International Affairs* 30 (3):401-403.
- Schmitt, Michael N. 2012. “Autonomous Weapon Systems and International Humanitarian Law: A Reply to the Critics,” Harvard National Security Journal online.
- Waldron, Jeremy. “Special Ties and Natural Duties,” *Philosophy and Public Affairs* 22.1 (1993)
- Williams, Bernard. 1993. “Moral Luck,” in *Moral Luck*, Daniel Statman (Ed.). State University of New York Press, Albany, New York, 1993, pp. 35-55.