Ethics and Information Technology (2019) 21:281–293
https://doi.org/10.1007/s10676-018-9472-6

ORIGINAL PAPER

Just research into killer robots

Patrick Taylor Smith[1]

Published online: 23 July 2018

Abstract
This paper argues that it is permissible for computer scientists and engineers—working with advanced militaries that are making good faith efforts to follow the laws of war—to engage in the research and development of lethal autonomous weapons systems (LAWS). Research and development into a new weapons system is permissible if and only if the new weapons system can plausibly generate a superior risk profile for all morally relevant classes and it is not intrinsically wrong. The paper then suggests that these conditions are satisfied by at least some potential LAWS development programs. More specifically, since LAWS will lead to greater force protection, warfighters are free to become more risk-acceptant in protecting civilian lives and property. Further, various malicious motivations that lead to war crimes will not apply to LAWS or will apply to no greater extent than with human warfighters. Finally, intrinsic objections—such as the claims that LAWS violate human dignity or that it creates 'responsibility gaps'—are rejected on the basis that they rely upon implausibly idealized and atomized understandings of human decision-making in combat.

Keywords Military ethics · Lethal autonomous weapon systems · Ethics and information technology

One Esk will shoot me if you order it. Without hesitation. But One Esk would never beat me or humiliate me, or rape me, for no purpose but to show its power over me, or to satisfy some sick amusement… The soldiers of Justice of Ente did all of those things.
-Ancillary Justice, Ann Leckie

Introduction

It would be an oversimplification to claim that we will soon see the deployment of lethal autonomous weapons systems (LAWS). Rather, it would be more accurate to say that weapons systems have long included autonomous elements.[1] For example, the Phalanx[2] close-in weapon system (CIWS), when activated, uses its own algorithm to identify and fire upon targets. The usefulness of the Phalanx is highly constrained, but other LAWS are being proposed, developed, and tested for a wide variety of capabilities and contexts. We should not be surprised if unmanned aerial, marine, and terrestrial vehicles—operating relatively independent of human control and guided by sophisticated and intelligent computer systems—begin to replace some of the jobs performed by human combatants in the relatively near future. This possibility has generated significant pushback from global civil society. These groups—represented most prominently by the Campaign to Stop Killer Robots[3]—advocate for a global governance regime that enforces a moratorium on the research, development, and deployment of LAWS. The goal of this paper is to evaluate whether cyberneticists, computer programmers, and other information technology experts and professionals are justified in pursuing LAWS research in

✉ Patrick Taylor Smith
polspt@nus.edu.sg

[1] Department of Philosophy, University of Twente, Enschede, Netherlands

[1] Devices that decide whether and when to explode, in the broadest possible sense, have been around since the nineteenth century in the form of naval mines, while homing missiles and proximity tracking explosives have existed since WWII. Before that, military use of trained animals—who have their own will and attack on their own volition in some cases—is almost as old as warfare itself. For a discussion of dogs in particular, see https://thestrategybridge.org/the-bridge/2017/12/9/autonomous-weapons-mans-best-friend.

[2] For a thorough description, see https://www.defenseindustrydaily.com/phalanx-ciws-the-last-defense-on-ship-and-ashore-02620/ (accessed 21-1-2017).

[3] See http://www.stopkillerrobots.org/ as well as a large scale petition signed by many information technology experts demanding a ban on LAWS research (https://futureoflife.org/open-letter-autonomous-weapons/).

spite of this skepticism. I will argue that LAWS research ought to continue as long as it is plausible that LAWS can perform warfighting tasks better than the good faith efforts of institutions populated by human soldiers, airmen, sailors, and marines (henceforth: human warfighters) and will then suggest that this state of affairs obtains now.

Autonomy and weapon systems

Since the Campaign to Stop Killer Robots does not intend to ban weapons that are currently in widespread use where there already exists significant automation, such as the Iron Dome, the Phalanx, and homing missiles, their target must be a system that possesses more than mere automaticity. Following Horowitz and Sharre (2015), I suggest that 'Killer Robots' then have the following features:

1) LAWS will be able to move about independently in the battlespace or, failing that, be able to intervene across geographically dispersed and contextually variable scenarios.
2) LAWS will have control over all of the functions necessary to engage targets and use lethal force: sensors, weapons fire, and follow up.[4]
3) Beyond being able to shut down the entire system, human beings will have limited ability to intervene and stop the decision-making process without paying high costs in terms of combat effectiveness or expense.

An autonomous weapons system, then, is one that can independently roam about the battlespace, use its sensors to acquire targets, and attack them where it would it be difficult (or exceedingly costly) for humans to intervene directly into the system's decision-making process. None of this requires general AI—where we might be compelled to grant the systems moral status and consider holding them morally responsible—and this autonomy is fully consistent with human beings having final control concerning whether and how the weapons are deployed and used. The question then becomes, "What are the relevant normative considerations that apply when deciding whether to research, produce, or deploy weapons systems that contain these capabilities?"[5]

Justice and LAWS research and development

Contrary to the demand for a moratorium by some members of global civil society,[6] I will submit that some are permitted to engage in further research of LAWS. However, justifying research often is quite different from justifying the full-fledged production and deployment of a weapons system. First, one might engage in research in order to *debunk* the idea that a system will have certain benefits. Second, justifying the production or deployment of LAWS would likely need to reference the opportunity costs that might attend the decision to use funding for LAWS development or deployment or the specific costs that might attend any particular deployment, which might have context specific factors that make LAWS deployment impermissible. So, even if LAWS could be justified as a research program, it might be true that we would not be justified in large-scale production and deployment. And this is broadly true of other kinds of research. The standard for determining whether a particular information scientist or roboticist *may* pursue the relevant research seems lower than the standard for determining whether a particular a system is actually used; after all, actual use generates real costs, costs which can—at least in theory—be avoided after we know more about the technology. What's more, we generally accept that researchers are entitled to wide latitude and discretion in their research choices. A person working on Alzheimer's or ALS does not do something immoral because work on malaria would more efficiently improve the welfare of potential sick people.

Yet, there must be limits to this discretion. It can be difficult to cancel programs once they have begun, and a research program could be so vicious or pointless that it would be wrong to pursue it. So, I want to suggest a *bootstrapping* account of when it is acceptable to research new weapons. That is, the justification of research ultimately depends upon the claim that, under plausible circumstances, the deployment of the technology will be an improvement in the relevant values over existing practices. In what follows, I will argue that we have good reasons to believe that advanced militaries operating in good faith[7] could use LAWS to improve their performance at

[4] The Campaign to Stop Killer Robots only calls for the ban on research of 'fully' autonomous weapons systems, which excludes the Phalanx and other systems in widespread use.

[5] For this reason, this paper will not be concerned with scenarios that involve robot uprisings as described in movies like *The Terminator* or books like *Robopocalypse*. Such scenarios are implausible and, even if we thought such possibilities were realistic, it is unlikely that a ban on researching LAWS would be relevant to preventing the 'x-risk' of human extinction or enslavement by robot overlord. First, AI research will continue outside the military context and, second, no one is arguing for bans on technology—military or otherwise—that can be controlled through a computer network and be used to kill humans. So, the ban will neither prevent the rise of a true AI nor will it prevent that AI from having access to weapons.

[6] The Campaign to Stop Killer Robots: https://www.stopkillerrobots.org/the-problem/.

[7] If one judged that one's military was not engaged in a good faith effort to fight just wars and fight them justly, then presumably *any* research program would be deeply suspect.

fighting wars justly. In doing so, we will need to examine both the *risks* associated with, and the intrinsic objections to, LAWS deployment.

Before I begin the analysis, I want to first present a set of considerations that should inform any answer about whether LAWS is too risky or too inherently wrongful to research. First, one should not fall prey to the fallacy identified by Mill in comparisons between socialism and capitalism: comparing the ideal operation of one system against the non-ideal operation of another (Mill 2008). It is, of course, true that LAWS will make mistakes that will lead to violations of the rules of just war, but so will human warfighters. One should be careful to ensure that one is comparing systems at fair levels of idealization.[8] For the purposes of this paper, I am going to compare how LAWS and human warfighters do and will operate in an advanced military that is committed to making good faith efforts to follow international humanitarian law that nonetheless will include some agents who will—either through negligence or choice—engage in morally prohibited uses of military force. Broadly speaking, I will be operating under the (stipulative and defeasible) assumption that the doctrine and policies of at least some advanced militaries are, if non-ideal, at least morally adequate. Consequently, research and development of modern weapons systems more generally is not inherently wrong for these militaries.[9] Nonetheless, it is possible that *specific* weapons systems—like LAWS—are intrinsically immoral, and I will discuss this possibility a bit later on. So, one of the key issues, then, is whether the inclusion of LAWS will improve on the moral performance of advanced militaries that are engaged in a good faith attempt to follow the principles of just war. Of course, many people might be skeptical of LAWS because they exacerbate what is already a morally prohibited activity, based upon a practical (contemporary states could never satisfy the requirements of just war) or absolute (all war is morally prohibited) pacifism. Such objections are important, but they are beyond the scope of this paper, which is interested in exploring whether we have *specific* reasons to reject LAWS over standard warfighting doctrines and strategies.

Second, we should attend to the distribution of risk as well as its overall level.[10] Most justifications of LAWS research and potential deployment depend upon the idea that it diminishes overall risk (Singer 2009, p. 398), but those arguments are importantly incomplete. That is, it would be unacceptable if we improved the aggregate level of risk to civilians and warfighters while imposing a disproportionate or unfair risk on a particular subset of the population. For example, it is plausible that a reduction in risk for warfighters and male civilians that required an increase in risk for children would be unacceptable even if the policy resulted in a net overall reduction. On the other hand, the dynamics of military actions are chaotic; any policy will have unintended and unforeseeable consequences as individuals or computer systems make decisions and take actions under the time pressure and information poverty of combat. So, there can be no guarantee that any military policy will be an improvement—in terms of risk of rights violations—for all individuals. What can be safeguarded against at the highest level of military doctrine and procurement, however, is whether that risk is imposed arbitrarily or disproportionately upon morally relevant classes of individuals. So, a change in military doctrine or equipment may impose a risk on a particular child in virtue of some complicated interaction of factors, but as long as that imposition is not reasonably foreseeable and the change will improve the risk profile of children in general (assuming children are a morally relevant category) in comparison to the relevant baseline, then it may nevertheless be justified.

From these general points, I suggest that we can derive an action-guiding principle for researchers that will be developing the algorithms and systems for LAWS:

> Research and development of LAWS is permissible if there is a reasonable expectation—given the limitations of computer engineering safety-proofing—that they will be deployed by the relevant advanced military in a way that generates a pareto improvement viz the risk profiles of all morally relevant classes of affected individuals over existing policy AND no intrinsic objections to LAWS deployment obtain.

[8] Consider the worry that LAWS will be hacked, as in the 2017 Open Letter (https://futureoflife.org/autonomous-weapons-open-letter-2017/). It is surely true that LAWS will be vulnerable to hacking. But human soldiers are vulnerable to similar issues. Undercover militants joining the police are a significant problem in counter-insurgency warfare, as are betrayals due to ideological sympathy, blackmail, or bribery. Setting intrinsic objections to the side temporarily, the question is not, "Can LAWS be hacked?" but "Will replacing or supplementing this particular job, post, or position with LAWS increase the risk of harmful behavior, given equivalent ideal and non-ideal assumptions?"

[9] Again, if we reached a judgment that a particular advanced military was indifferent to the principle of just war, then that would give us good reasons not to continue LAWS research for that particular military.

[10] In setting up the issue this way, I am deeply indebted to Simpson and Mullers (2016). They argue that we should understand just war principles in terms of fair distributions of risk. My paper is different in at least three ways. It sets out the dynamics of the modern military system in order to more fully describe the appropriate baseline, it uses that baseline to undermine intrinsic objections, and it is about the morality of *research* rather than the morality of deployment.

A few clarifications are in order before we proceed. First, whether LAWS research is permissible could vary with developments in engineering or computer technology. It may turn out that developing information systems, sensors, and the other technologies to allow for adequate targeting judgments will be much harder or much easier than we anticipate. So, the permissibility of LAWS research and development may change over time. Second, the 'reasonable expectation' of the relevant improvement on current military practices will need to incorporate a basic presumption that weapons systems predicated on information technology will have greater intrinsic risk due to the limitations of computer engineering (Brooks 1987). Third, the 'pareto improvement' requirement is that at least one morally relevant class of individuals will see their risk profile improve and that no other classes are made worse; LAWS need not improve the risk profile of all classes. It need only be the case that it makes no other class beside the improved group worse off than before. So, I claim that if I show that no intrinsic objections apply to LAWS and that LAWS likely represent a pareto improvement over current practice, then it is permissible to engage in LAWS research and development for militaries that are sincerely engaged in an effort to improve compliance with just war principles.

Just war theory, international humanitarian law, and killer robots

In order to evaluate whether research into LAWS should continue, we need to understand whether the *deployment* of LAWS could be justified under sufficiently likely circumstances. A common and powerful set of objections to the deployment of LAWS is that they will be inferior to human soldiers at following the principles of just war—including and especially the *jus in bello* requirements of distinction and proportionality—and their political manifestations in international humanitarian law and the law of armed conflict (ICRC 2016).[11] It is important to separate these objections into two categories: intrinsic and contingent. Intrinsic objections are those where the essential nature of LAWS makes it *necessarily* true that they violate some moral principle. For example, one might worry that being killed by a machine is inherently or intrinsically degrading (Asaro 2012). The idea is that appropriately complying with *jus in bello* requires human agency and that killer robots must, by their very nature, fail to satisfy that requirement. One reason this might be true is that removing human beings from the decision-making loop makes it impossible to determine who ought to be held responsible when the machine makes a mistake and violates the rules of war (Sparrow 2007). The fact that a machine's decisions—represented by a complex program that may interact in chaotic and unpredictable ways with situations that are impossible to fully specify in advance—mediates between any human's decision and the final result makes attributions of responsibility difficult. On this view, it is inherently disrespectful, undignified, or otherwise wrongful to use weapons against individuals in a way where no one can be held accountable.

Contingent objections, on the other hand, suggest only that it is more likely that LAWS will violate just war principles and that, consequently, it will be excessively risky or even negligent to engage in LAWS research and development (ICRC 2016; HRW 2016). One such source of risk is that LAWSs will be indiscriminate: they will target non-combatants deliberately or accidentally. There are two sources to this worry, one technical and one political. On the technical side, the worry is that it will be difficult or impossible to program the systems such that they will be able to make the complicated judgments concerning who is a civilian and who is a combatant, which might depend on ascertaining intention as well as picking out the precise, salient details in a complex environment. On the political side, the concern is that the conscience of individual soldiers represents an additional check on commands to attack civilians that LAWS will lack. On the contrary, LAWS will simply follow the instructions they are given and are unable to resist alterations in their programming. So, if policy-makers or commanders issue an illegal order, the LAWS will not be in a position to resist the command. Of course, we might attempt to program the laws of war into the LAWS, but this may not prevent bad actors from changing the program and thus allowing the LAWS to engage in atrocities. Similar worries apply to proportionality judgments. Determining whether a LAWS should direct fire at an insurgent sniper in an urban home requires difficult judgments about intentionality, relative value, and an understanding of the context that will likely elude even sophisticated autonomous systems. One, again, might worry that computers will be unable to reliably make good judgments even in contexts where efforts are made in good faith or that LAWS will not be a check on bad actors who are indifferent to the laws of war. Finally, problems arise at the intersection of these principles. The computer program would need to be immensely complicated to guide LAWS through these judgments and that complexity would undermine our ability to predict how the systems will behave in situations where the various judgments are all operating together, where they must interact and even come into conflict.

[11] Ibid. For the canonical statement of just war principles, see Walzer's (2002). For international humanitarian law, see the Geneva Conventions, especially Article 51 of the First Protocol Relating to the Victims of International Armed Conflicts, which concerns discrimination and proportionality.

Additionally, one could argue that that the primary benefit of LAWS—immunizing human soldiers from harm (Boland 2007; DoD 2005)—leads to problems in the long run. If military operations become essentially costless to the attackers, then military force may no longer be understood as a last resort, potentially generating a kind of moral hazard. If the costs of a risky activity—military operations—are entirely externalized in that they are only suffered by targets, then we would expect attackers to use military operations more frequently. Some have argued that we have already seen this dynamic at play in the use of armed drones in American counter-terrorism activities. Drones incentivize overuse by making military operations possible when ordinarily they would be prohibitively expensive. As a consequence, we are experiencing an expansion of pseudo- and proxy wars with unclear objectives or endpoints. If LAWS produce additional military capabilities that are relatively costless to potential attackers, then we might expect this tendency to worsen.

The case for LAWS as pareto risk improvement

The argument of this section lies in the relationship between warfighter and civilian risk. One advantage of LAWS is that it reduces risk to soldiers, airmen, and marines in combat (Boland 2007; DoD 2005). So, the key issue is whether the improved position of warfighters is purchased at the cost of increased risk for others. I argue that not only will LAWS likely not be worse than current policy for civilians; they will likely be better. There are two mechanisms by which LAWS can improve on human warfighter performance in advanced militaries. First, LAWS will avoid intentional and unintentional violations that are inevitable with human warfighters given the limits of human agency. Second, since LAWS deployment will make force protection less morally or practically necessary, it makes it possible for militaries to adopt casualty-acceptant policies in order to reduce risks to the civilian population. Of course, these claims depend upon LAWS being able to achieve certain capabilities before deployment, and I leave it to the experts as to whether these can be achieved under a plausible time frame.

Let us begin with the ways in which LAWS are likely to avoid both intentional and unintentional deviations from just rules of engagement. Rules of engagement are the policies that regulate the deployment of force in advanced militaries and are typically determined by the officer corps with input from the political leadership; rules of engagement are context sensitive and can vary within a war, a campaign, or an operation. Just war violations by human warfighters typically occur either because they are guided by problematic rules of engagement or because they fail to follow just ones. In the latter case, human soldiers often fail to discriminate or engage in disproportionate uses of force because of anger, fatigue, or fear; LAWS experience none of these. So, while it is true that human soldiers will have more sophisticated faculties of judgment, it does not follow that the human soldiers will be better at following just rules of engagement than LAWS. In the former case where superiors assert problematic rules of engagement, we have little reason to think that human warfighters will be better at *rejecting* unjust rules of engagement. Humans typically follow orders and bad political actors usually have little difficulty in getting the rank and file of militaries to accept their commands. While it is indeed true that sometimes soldiers disobey orders in order to protect civilians and obey the laws of war, often at great personal risk, these actions are rare and unsystematic.[12] Furthermore, it may very well be true that many commands to commit atrocities are not given because the commanders judge that their soldiers would not follow them. Yet, there are negatives to human agency to balance against these positives because it is also true that warfighters disobey rules of engagement when those commands are to engage in greater protection of civilians. Warfighters also engage in mutiny or coups against legitimate democratic regimes out of tribalism or greed. Humans will be motivated to save civilians in ways that LAWS will not, but they will also be motivated to violate the laws of war in ways that LAWS will not.

Opponents of LAWS tend to ignore the potential wrongs and atrocities that can only be committed by human soldiers. LAWS will not degrade, rape, pillage, or kill for pleasure. Sexual assault is an unfortunate commonplace in warfare. Rape is often considered both a 'spoil' of war used to incentivize soldiers to fight and used deliberately as a weapon (Schwartz 1994; MacKinnon 1994). Sexual assault can be used to demoralize civilian populations and to force them to abandon support for insurgents as it was in Chechnya (Rousseva 2004). Or, conversely, it can be used as part of a deliberate, genocidal campaign to eliminate a hated ethnic group as it was in the former Yugoslavia (Barstow 2000). Generally speaking, war operates at the liminal space between legality and illegality, and this provides opportunities for human agents to engage in all sorts of criminality, domination, exploitation, and abuse. The greater the number of soldiers, the less adequate the supervision, and the greater their stress, the greater the likelihood that warfighters will

[12] The heroic actions of Hugh Thompson and his helicopter crew at the My Lai Massacre during the Vietnam War are instructive (Angers 2014). Upon observing a massacre of Vietnamese civilians by American soldiers, as ordered by Lieutenant Calley, they saved many innocent people by physically blocking soldiers from approaching parts of the village and then reporting the massacre to superiors. Yet, despite the fact that Lieutenant Calley was acting in direct contravention to his orders, he had no difficulty getting the vast majority of the soldiers under his command to obey. What's more, it is worth noting that Thompson and his crew were *not* under Calley's command.

abuse their position. The use of LAWS can have a spiraling positive effect on these dynamics where fewer warfighters leads to greater and more effective oversight over the human warfighters that remain in theater, while less stress and fear for those in combat all jointly lead to fewer just war violations. So, from the standpoint of risk, human warfighters would need to be *considerably superior* to LAWS in some other area to compensate for the specific violations that can only be committed by humans.

Even if we grant sufficiently superior performance to human warfighters for a particular set of rules of engagement, we should not hold the rules of engagement constant. Human warfighters will—due to military necessity—need to adopt more civilian-risk-acceptant military doctrine and tactics than LAWS. Many critics of 'killer robots' misunderstand the source of most civilian casualties in modern warfare. Beginning in the middle of the nineteenth century, a series of ideological, technological, and organizational innovations generated what has come to be called the "modern military system."[13] Automatic weapons, explosive weapons, and rapid fire, breech-loading artillery have created an 'empty battleground' whereby any target that can be easily seen can be destroyed. The open field maneuvers of the Napoleonic Age and the early parts of the American Civil War became suicidal. This gave rise to fire and movement tactics by which defensive positions needed to be suppressed by indirect fire or by automatic weapons before one could advance. Cover and concealment becomes essential on both offense and defense; for the offense, the goal is to take advantage of 'dead zones' that are not covered by the automatic weapons of the defender and, for the defender, cover and concealment reduces the effectiveness of suppressive fire.[14] Most weapons fire in the modern system is suppressive; offenders need it to advance and defenders need it to pin and fix the offenders for their own indirect artillery. As a consequence, individual warfighters usually only have a vague sense of who and what they are attacking. It is usually impossible—on the modern battlefield—for a particular soldier to observe their enemy for longer than a moment; they will be dead long before they can. These fire and movement tactics are the foundation of combat, from low-intensity counter-insurgency operations to full-spectrum warfare between conventional armies. In other words, in the time it takes to make the individual judgments that most philosophical accounts of just war assume (as well as most critics of LAWS), the warfighters will likely be dead. This is why rules of engagement are so essential: they set the parameters of action *in advance* and serve as constraints on the judgements of individuals in combat.

There are two important corollaries to the modern system. First, the level of civilian casualties is frequently determined by how discriminate indirect and suppressive fire can be made to be. This is well illustrated by an analysis of Iraqi civilian casualties during the height of the Sunni insurgency after the American invasion in 2003.[15] Civilian casualties were driven by two primary factors. First, there was the failure of the American and Iraqi governments to police and prevent ethnic cleansing in Baghdad. Second, civilian deaths were caused by indiscriminate suppressive fire in battles between coalition forces and the insurgents. Broadly speaking, then, the majority of civilian casualties in modern warfare result from two distinct modalities, usually in the context of a civil war or insurgency: terror-motivated deliberate targeting outside of combat and indirect fire within it, driven by the demands of the modern military system.

It is worth considering whether LAWS will exacerbate the intentional modality for two reasons. First, LAWS will likely never say 'no' and refuse to use force when ordered to do so by a person with the appropriate authority. Yet, this does not represent a significant increase in risk when compared to human warfighters. Two pieces of evidence support this claim. First, insurgent or state forces that engage in terror campaigns have not, at least not obviously, failed because the human warfighters that make up those forces have moral objections to terror tactics. The Sunni insurgents, Milosevic's military, or Putin's forces in Chechnya succeeded or failed due to other factors; in fact, these groups demonstrated remarkable resilience in the face of incentives to defect. Of course, both insurgents and state forces can, using terror tactics, lose legitimacy and support amongst the civilian population. This may lead to their defeat, but it is hard to pinpoint examples where moral atrocities by a military force lead to operational defeat due to defection or resistance within that force itself. Second, our most sophisticated theories of civil resistance show that protests and civil disobedience are most likely to succeed when military and paramilitary forces refuse to suppress them. This seems to support worries about LAWS because they will always follow orders to suppress civil resistance. Yet, the failure to suppress is most closely correlated with *elite* division, not division within the rank and file members of the forces (Chenoweth and Stephan 2011). That is, military forces will refuse to quell a

[13] My account of the 'modern military system' is based on the analysis of Chap. 3 of Biddle (2006). I've also benefitted from Pollack (2002), especially Chap. 1. For a good illustration of these concepts in military practice: see Department of the Army (1998).

[14] Swinton (1986) is a commonly used illustration of these principles in military academies.

[15] According to the Iraqi Body Count, only 13% of civilian casualties in Iraq could be attributed to coalition direct fire; the majority of casualties were the result of indirect combat between insurgents and coalition forces or committed by insurgent forces alone: https://www.theguardian.com/news/datablog/2012/jan/03/iraq-body-count-report-data.

crowd when there is ambivalence at the top of the command chain about whether doing so is a good idea; spontaneous refusal by enlisted personnel, NCOs, or low-ranking officers is exceedingly rare and usually insignificant in operational terms. In other words, if you want a military to act justly, then train them to act justly, order soldiers to follow just rules of engagement, and develop a professionalized officer corps with good civil-military relations. It is, however, highly unreliable to expect individual agents—especially those who are strongly socialized towards obedience and are faced with considerable risk and uncertainty—to routinely act against their comrades and their orders. And even in the heroic and rare cases where it does occur, resistance is rarely systematic. So, once we include the various ways that LAWS could improve physical security, it seems unlikely that deliberate terror campaigns will be affected by LAWS deployment other than making them more difficult to foment.

Turning to the indirect fire modality, the second corollary of modern military systems is that *force protection* is the clear driver of permissive rules of engagement and negligent failures to obey more restrictive ones. The main reason that we need to accept some degree of suppressive fire is to avoid military casualties in order to accomplish the relevant military objectives. This is not to say that we need to accept *all* suppressive fire; some—perhaps even most—rules of engagement are overly permissive from the standpoint of just war theory. Yet, the bare fact that modern militaries are going to endorse rules of engagement that allow soldiers to fire somewhat indirectly and blindly is not going to change as long as the modern system is still in place and soldiers need to protect themselves. What's more, insofar as these tactics are necessary for basic military effectiveness, then they can be justified—if the civilian casualties are shown to be proportionate—under just war theory and international humanitarian law. These tendencies are exacerbated in counter-insurgency warfare. Consider the following case, common in Iraq and Afghanistan:

> CAR BOMB: A convoy is travelling along a major arterial in an urban center. A squad has closed an intersection ahead of the convoy, setting signs and barriers indicating that all cars must stop 100 feet short of the intersection. A car approaches and then passes the boundary with no indication of stopping. The 0.50 cal gunner needs to choose whether to open fire.[16]

Is the driver an innocent civilian who did not notice the signs or a suicide bomber intent on attacking the convoy? It is hard to see why, *in principle*, a human being will be better at making this decision than a LAWS. The human will undoubtedly have some advantages, an empathetic ability to put themselves in the position of the driver, but those advantages may require some time to function properly. The LAWS will have advantages as well. After all, a LAWS will likely have better sensors, a greater ability to integrate them, and a faster reaction time. For example, perhaps the LAWS will be able to integrate distinctive noises, infrared sensors, and immediate intelligence updates to make a judgment about the likely status of the truck that is more likely to be correct than a human. But the more significant point, for our purposes here, is that whatever temptation we have for allowing the 0.50 cal gunner to operate according to rules of engagement that permit him to fire once the truck crosses the boundary is based on force protection. The gunner is responsible for the lives of the convoy and his squad. And why not require the squad to physically stop the truck rather than firing upon it? The reason is the same: the truck might be a suicide bomber and the squad has a legitimate interest in their own protection, either intrinsically or because a dead squad will fail to generate morally relevant military advantages.[17]

So, force protection generates the justification—both pragmatic and moral—for adopting tactics, doctrine, and rules of engagement that place substantial risks upon civilian populations. Yet, LAWS do not have the same force protection considerations; they are replaceable machinery and not human beings.[18] This has two interlocking and mutually reinforcing benefits. First, the lack of a human occupant allows the LAWS to take risks to safeguard the interests of civilians that we would—in other contexts—take to be unacceptable to human occupants, pilots, and soldiers. Second, the biological and psychological nature of human beings set constraints on the capabilities of vehicles, sensors, and communication systems. Without the need to keep a human being alive and conscious, we can design weapons systems with capabilities that may very well make them less likely

[16] A veteran and student in my "Ethics of War and Peace" course suggested this case based on his experience in Fallujah.

[17] There is considerable controversy as to *why* force protection is morally important. Walzer (2006) argues that force protection is only relevant due to military necessity: dead soldiers cannot achieve the military objective of winning a just war. Others (Kasher and Yadlin 2005), however, have argued that soldiers need not take on risks to themselves to spare civilian lives. I remain agnostic; my paper only requires that military behavior is driven by force protection and that this is at least sometimes justified.

[18] Singer (2009) makes the point that LAWS can be more 'conservative' than human warfighters. My view moves beyond his in several ways. First, it bases the claim on an understanding of the modern military system. Second, it offers the normative basis for the civilian-risk-acceptant behavior of human warfighters. Third, it describes the knock-on dynamics of this advantage for LAWS. Fourth, it uses these insights to engage with intrinsic objections as well. Fifth, Singer does not describe the issue in terms of the fair distribution of risk and so does not offer adequate normative guidance on the question.

to engage in the kinds of firefights that get civilians and soldiers killed. We can already see this in the use of unmanned aerial vehicles. Drones can loiter over the battlespace for extended periods of time, placing ordnance far more precisely than a human piloted fighter-bomber which must fly faster and higher to avoid attack. Drones can do this for two technical reasons: the lack of a human pilot means that space and equipment that are usually dedicated to keeping a pilot alive can be dedicated to increasing longevity, and the length of deployment need not be constrained by pilot fatigue. However, there is a *normative* reason why drones can fly low and slow: it does not matter that much if they are shot down. And there is good evidence that drone attacks are much less damaging to civilians than corresponding attacks by fighter-bombers, artillery, or cruise missiles.[19] So, the force protection potentialities that LAWS represent can contribute to civilian protection by eliminating the primary motivation for engaging in military behavior that risks civilian casualties. This allows advanced militaries to endorse and follow rules of engagement that are risk-acceptant with regards to the LAWS but risk-avoidant when it comes to civilian life.

This leaves us with the moral hazard argument. Robert E. Lee once said, "It is well that war is terrible, lest we grow too fond of it."[20] These arguments assume a similar worldview: the costs of war are what fundamentally constrain policymakers from engaging in it. If we lower the costs of making war, then we will fight more wars. Drones are taken to be a specific example of this trend: the United States has engaged in a policy of semi-permanent warfare against an amorphous, transnational entity, but it does so in ways that are comparatively costless from its perspective. As a consequence, American leaders have never been held to account and the war drags on. So, while drones may be superior to other military options for particular missions, the availability of drones as a tactic leads to many more attacks, including and especially those attacks that are disproportional in the face of marginal military benefits.

This argument is, at best, speculative. First, while there is some surface plausibility to the idea that lowering the costs of war will increase its incidence, neither a close study of history nor polling data support it.[21] There is no clear correlation between the public's support—either in terms of a base level or marginal change—for a war and expected or actual military casualties; judgments about the value of the military intervention dominate. Furthermore, incidences of war, military casualties, and civilian deaths from war all have been steadily decreasing for the past few decades, including after the introduction of drones (HSP 2013). That is, while war has become less costly in terms of military casualties, we have seen less war and fewer civilian deaths. Three general tendencies might explain this. First, greater military casualties might create a path-dependent dynamic; politicians need soldiers' deaths to 'mean something' and so they continue with failed policies. Similarly, if it is expected that military casualties will be high, then one might expect and demand commensurate benefits for those costs. The other side will anticipate these demands, making them less likely to compromise, which the initial side will also anticipate. In other words, an expectation that one will suffer more severe casualties might make it harder to go to war but might also make wars more intractable once they start. So, we might expect more wars under a LAWS regime, but this does not obviously show that human welfare will be net worse. Third, the expectation of higher casualties might generate more permissive rules of engagement, leading wars to be far deadlier even if they are less common. Historical examples bear this out. The much higher potential costs of intervening in Cambodia and Laos did not stop the Nixon Administration from expanding the geographical scope of the Vietnam War. Instead, the administration adopted more permissive rules of engagement and less discriminate uses of airpower. What's more, the majority of US casualties occurred after Nixon was elected on a platform the end the war. This was driven, in part, by the need to achieve 'peace with honor' and to receive something in return for the expenditure of American blood and treasure. Similarly, it is unclear why the alternative to drones are not simply more deadly and dangerous attacks of other kinds. Perhaps the Pakistani military would have engaged in a correspondingly greater number of search and destroy missions in Waziristan without the drone attacks, and those missions kill far more civilians on average than drones. Finally, as the previous example suggests, it is not obvious that the 'marginal' missions that will or will not be performed with the costlier tactics will correspond particularly well to the set of unjustified

[19] This article aptly summarizes the available evidence that drones are more discriminate than other tactics: http://www.slate.com/articles/news_and_politics/foreigners/2015/04/u_s_drone_strikes_civilian_casualties_would_be_much_higher_without_them.html. A few other studies purport to show that drone attacks are not more accurate, but they often rely on problematic comparisons. For example, some studies show that fighter-bomber attacks on ISIS military positions kill fewer civilians per sortie than drone assaults in Pakistan or Yemen. But this is not an apples to apples comparison since ISIS is engaged in conventional, full spectrum operations with operational lines and distinct warfighters. Attacks in Yemen and Pakistan are aimed at insurgents who deliberately hide amongst civilians. Conventional tactics by the Pakistani military in these areas generate civilian casualties that are orders of magnitude higher than drone attacks.

[20] As reported by Edward Porter Alexander (1907).

[21] For example, neither Larson (1996) nor Gartner and Segura (1998) assert a straightforward relationship between casualties and support for a war. First, war support always erodes over time but it can be revived through victories even in the face of high casualties; the way the war is fought and perceived plays a key role. Second, casualties cannot explain the *level* of support for war even in those cases where casualties can explain a reduction in war support.

missions. After all, policymakers are willing to accept costs in relation to perceived benefits. Yet, if those policymakers have skewed perceptions (e.g., the marginal reduction of the risk of a terrorist attack is far more important than the prevention of a genocide in a strategically irrelevant country), then the military attacks that are eliminated for being 'too costly' might be precisely those attacks that could very well be justified. In other words, military interventions that serve 'national interests' will likely be acceptable to both civilians and policymakers while military interventions that serve the common good might fail the test.

So, the argument that a researcher could reasonably conclude that LAWS—if it developed particular capabilities—would represent a pareto improvement over current policy seems complete. Vulnerable classes of society will not be subject to risk generated by human desire for degradation, domination, and cruelty. Nor will civilians be subject to mistakes or weakness of will generated by fatigue, anger or fear. LAWS will not be significantly more likely to obey illegal orders than human warfighters. And most importantly, LAWS will make possible more restrictive rules of engagement that will allow advanced militaries to prioritize civilian interests over force protection. For these reasons, a current researcher who wished to pursue LAWS research need not refrain from doing so because of contingent objections to LAWS.

Intrinsic objections

However, the LAWS researcher is only permitted to continue her work if both contingent and intrinsic objections fail. The previous section was a response to the former, but we still need to consider the latter. There are two key intrinsic objections. First, Robert Sparrow (2007) has argued that there is a 'responsibility gap' in LAWS deployment. If a LAWS fails and violates the laws of armed conflict, Sparrow argues that there may be no agent who can be held accountable. He then suggests that this represents a necessary violation of just war principles. Thus, it is inappropriate to research, develop, or deploy weapons systems where no agent can be held accountable for their violations of just war principles. Second, some (Asaro 2012; Johnson and Axinn 2013) have argued that LAWS violate human dignity. The crux of these arguments is that the decision to end a human life must be—in some meaningful and important sense—made by a human being. So, presumably, one is not permitted to engage in research that necessarily undermines human dignity. I respond to both objections in the same way: neither amounts to a problem for LAWS *in particular*. That is, while there are interpretations of their claims that would apply to LAWS, they are tremendously over-inclusive in condemning practices that have historically been considered acceptable and where their permissibility is fairly intuitive.

Let us begin with Sparrow's argument concerning the 'responsibility gap.' Suppose that a computer system makes a decision to deploy lethal force in a way that is unjustified where such failures are generally but not specifically foreseeable. That is, the chaotic interaction of environment, rules of engagement, and the system's programming will predictably generate failures that are themselves individually unpredictable. When this happens, Sparrow argues that there are three possible agents we might hold accountable for the failure: the programmer, the commander, or LAWS itself (Sparrow 2007, pp. 69–73). None represent attractive options. The programmer is too distant, and the LAWS violation will likely be a consequence of the interaction of factors—such as mission and environment—that are beyond the programmer's control. So, attributing moral and legal responsibility would be unfair. The commander, on the other hand, is not responsible for the programming and it seems possible that the chaotic interactions of various elements could lead the LAWS to fail in ways the commander could not reasonably be expected to foresee. If the commander took appropriate precautions, it seems inappropriate to blame her for the failure. The final option, holding the LAWS responsible, seems the least plausible. After all, we are not discussing a system with general AI that might have interests we could affect. No punishment or incentive structure could be put in place to ensure that these autonomous systems followed the relevant laws. Even if we wanted to hold these systems accountable, it is not clear that we could. So, there are cases where a LAWS failure would generate a rights violation but no correlating accountability for that violation, which is contrary to the requirements of just war.

I object to this argument in two steps. First, I want to suggest that we need to be clear on why human accountability is necessary for full compliance with just war principles. There are, it seems, two possibilities: human accountability for just war violations is either causally essential to any just war regime or it is an intrinsic, moral requirement of just war principles themselves. That is, is accountability central to just war theory because it serves a key *policing function* that incentivizes actors to follow its principles or, is accountability important because human beings who have been subject to military force *deserve* to be able to hold someone accountable? If it is the former, then the lack of accountability is merely one more risk factor amongst many and its importance would depend on whether that policing function played an important role in just behavior by warfighters in modern militaries. If it is the latter, then Sparrow is committed to a strong version of the correlation thesis where every rights claim must have a correlating duty such that some agent has failed in their obligation if that right has been violated. That is, if accountability is morally necessary for just war compliance, then it seems like any military system that generated responsibility gaps would violate this requirement.

Thus, Sparrow is committed to either a controversial empirical claim or a controversial normative claim.

The second step in the argument is to show that existing military practices—including practices that are widely viewed as generally permissible—undermine the effectiveness of the causal claim and would be ruled impermissible by the normative claim. Let's recall some key features of the modern military system. Fire and movement tactics require the coordination of many tactical elements, most of whom do not have direct sight on the battlefield. Furthermore, individual warfighters are ill-equipped to make determinative proportionality judgments, and people away from the battlefield are ill-positioned to make determinations about the location, necessity, and effectiveness of indirect fire. As a consequence, warfighters need to rely on well-considered and justified rules of engagement in making judgments about discrimination and proportionality. In this context, neither the empirical nor the normative claim looks especially plausible. The empirical claim suggests that the system of military justice plays a key policing role, ensuring that soldiers—in combat—follow the principles of just war. There are several problems with this. First, military justice proceedings for in-the-moment judgments about use of fire are very rare and almost always involve soldiers who deliberately fire upon civilians or warfighters that deliberately disobey the rules of engagement when deploying fire negligently (Sikkink 2011). Within the context of the rules of engagement, military justice proceedings appropriately give wide latitude to the forces on the ground. Second, most negligent fire that wrongly risks civilian life is performed by warfighters that believe they are defending their own lives; it seems implausible—if the soldiers are not themselves predisposed to take the relevant risks—that they will be moved by the threat of punishment if they already fear for their lives. Military justice plays an essential role in maintaining discipline and ensuring that doctrine and the rules of engagement are followed, but it is a weak reed for ensuring that soldiers follow the principles of just war in the heat of combat and in the face of difficult proportionality and discrimination judgments; training, doctrine, and institutional culture are much more important.

Turning to the normative claim that accountability is a moral requirement of complying with just war principles (Sparrow 2007, p. 67), the modern military system creates a 'responsibility gap' for human warfighters as well. Let's return to the case of CAR BOMB. Let us suppose that the commanding officer—deeply fatigued from planning the mission because enemy mortar shelling prevented her from falling asleep—drafts an operations plan that is problematically vague about the rules of engagement, with one plausible interpretation imposing marginal and likely unjustified risks on civilians heading towards the checkpoint. Further, let's suppose that a typical warfighter would likely not have misinterpreted the ambiguity in a civilian-risk-acceptant way but that the 0.50 cal gunner was angry and distracted because two of her platoon mates were injured in an IED ambush earlier in the week. As a consequence, she interpreted the orders in ways that were excessively lenient concerning the use of force. She fires too soon, killing or injuring the civilians in the car unjustifiably. Who is responsible for this violation of just war principles? The commander with a full capacity subordinate would not have violated the rules of war, and a subordinate with a full capacity commander would not have either; they had to act in concert. What's more, they both have reasonable excuses for their individual failings that normally would have been harmless but for an unfortunate confluence of mistakes (including the poor judgment of the civilians driving towards the checkpoint). My view is that the civilians have been wronged and, as a consequence, the broader political system owes them reparations and compensation as a matter of justice. While there are a variety of ways in which that *system* failed those civilians, no particular person is responsible for the wrong, and no one ought to be punished. Of course, the commander and the subordinate both did their job poorly, but it would be unfair to hold them responsible for their mistakes in this case. It is important to see that this example does not show that it would be difficult—either epistemically or practically—to hold these individuals accountable. Rather, precisely like the commander or the programmer, CAR BOMB shows that there are violations of just war principles within the modern military system where it would be *unjust* to hold any particular agent accountable for the violation.

Of course, one might try to argue that the commander and the subordinate are both partially responsible and ought to be held accountable. Perhaps this is true, though it would not be difficult to construct a case where the 0.50 cal gunner bore essentially no responsibility for the mistake due to their reasonable decision to follow flawed rules of engagement. This would be much like how the LAWS would bear no responsibility for its mistake. Even if we grant that many agents can have partial responsibility for these kinds of mistakes, it is clearly true that a human warfighter makes complicated, split second judgments that intersect with environment, training, and orders in chaotic ways such that a 'responsibility gap' remains. And this goes far deeper than the specific rules of engagement for a particular mission. After all, perhaps the gunner must use this especially deadly kind of machine gun or ammunition because of appropriations decisions made by the defense-industrial complex years ago. Or perhaps she has not been trained to take out the tires or engine block, limiting her options at the moment of decision, because her military and civilian superiors decided—not without justification—that such training was usually unnecessary in full spectrum operations. They were unable to switch training easily or quickly in the face of the need to intervene in an

unplanned-for crisis. Modern militaries are large institutions that make training, procurement, and doctrine decisions years in advance and cannot change policy quickly. In other words, responsibility for a particular violation of just war principles can be diffused throughout the system with the trigger-puller being quite innocent. It is hard to see why this is so fundamentally different from the possibilities offered by LAWS. In other words, in order to condemn LAWS, we would need to condemn standard military practice driven by the requirements of the modern military system: both lead to responsibility gaps on a consistent basis.

Sparrow would like to resist this conclusion. In order to do so, he suggests that there is a distinction between technologies based upon the *likelihood* of a responsibility gap. He writes:

> However these accidents [where no one should be held responsible for the deaths in violation of the rules of the war] represent regrettable, if inevitable, failures to live up to principles of justice in war fighting. If the nature of the weapon, or other means of war fighting, is such that it is *typically* impossible to identify or hold individuals responsible for the casualties that it causes, then it is contrary to this important requirement for *jus in bello*. (Sparrow 2007, p. 67)

So, Sparrow is forced to retreat from the claim that a technology or way of fighting is necessarily wrong if there is no human accountability to the claim that these practices are wrong if it there is 'typically' no human accountability. Yet, his arguments about the programmer or the commander do not rely on the LAWS 'typically' being unaccountable, only that they may generate unaccountable casualties. For example, he argues there is a "possibility" (Sparrow 2007, p. 70) that an autonomous system may act beyond its programming, which is not a claim that such behavior is typical. Second, he argues that the use of LAWS generates a "risk" that autonomous weapons will act in ways that commanders neither commanded nor could have anticipated (Sparrow 2007, p. 71), but this again falls short of claiming that this behavior is typical. If my above arguments are correct, then these risks are also present with human warfighters. Sparrow seems to be arguing that LAWS will generate responsibility gaps beyond some threshold of risk while contemporary practice does not. Yet, it is hard to imagine a reason why the normative claim—that human accountability is morally necessary—should be indexed to a threshold, or why we would want to abandon superior laws of war performance in order to fall below the threshold of not 'typically' causing responsibility gaps versus only sometimes causing them. In other words, if Sparrow conceives of the requirement for human accountability as absolute, then essentially all modern warfighting doctrine and practice would be impermissible. Conversely, if he conceives of the requirement in a probabilistic fashion, then the argument is no longer intrinsic and we can return to the risk analysis in the earlier part of the paper.

Sparrow might suggest that it is morally necessary that we be able to *attribute* the act of killing to a human agent, even if we do not ultimately punish or condemn a person for it.[22] This seems to be Peter Asaro's position: he relies upon the principle that human dignity demands that the decision to kill another human being must be made by another human being. As a consequence, there must be meaningful human control within the decision chain to use deadly force. Since LAWS, ex hypothesi, do not do this, LAWS are intrinsic violations of human dignity. Of course, Asaro needs to defend more than the claim that it would be better for human dignity if all decisions to kill were made by humans; he needs to defend that idea that the threat to human dignity by computer decisions is sufficiently important that even real gains in terms of fewer people killed, raped, or degraded cannot compensate or override the dignity loss caused by LAWS.

I present two responses. First, the insights about the modern military system appear to decisively undermine it. Perhaps it is true, in some ideal sense, that aggression not only should be directed *at* its proper target but also *by* a rational agent making an autonomous and fully explicit judgment that this particular exercise of military power is required by the principles of just war. But real world combat is not especially accommodating; most exercises of military power are done according to previously ingrained training and following the rules of engagement of superior officers. And as we just saw, the pressures and constraints of the modern military system makes that kind of explicit, individual, case-by-case judgment extraordinarily difficult and even morally undesirable. It is unclear why the normal decision process of the typical warfighter making a decision to fire should be placed in such high esteem. While we might grant that the ideal is a worthwhile one, the distinction between LAWS and *current* practice is much smaller than Asaro is willing to admit. One can respond to this claim by arguing that current practice is unacceptable and essentially all military actions in the modern era are unjustified, but this is a substantial bullet to bite.

The second response is that we do not accept this sort of human dignity argument in other contexts where human life is at stake. Far more people have died in the United States in the last few years from traffic accidents than terrorism or combat. Yet, consider the humble traffic light.[23] In many cases, traffic lights are governed by complex

[22] On the (possibly problematic) difference between responsibility as attributability and responsibility as accountability, see Watson (1996) and Smith (2012).

[23] Asaro, Johnson, and Axinn do not seem to apply these arguments to self-driving cars, which also involve AI systems making life or death decisions.

algorithms that receive inputs from sensors in the road that lead these devices to make 'decisions' about traffic flow. These decisions—including and especially those rare cases where a 'mistake' is made—carry with it the risk that human beings will die. Of course, human agents—for now—will make various decisions in the context of the traffic light, but none of these agents are making explicit judgments to crash into each other. The clear fact of the matter is that a human decision-maker—somewhere in the chain—decided on the relevant risk tolerance: they compared the likelihood that a traffic light decision would result in an accident and decided that the net cost to traffic flow (perhaps the costs to the municipality making the procurement decision, or more problematically, to the bottom line of the traffic consulting or urban planning firm) that preventing the accident would require was not worth it. Similarly, a human being somewhere in the chain will decide on the relevant risk tolerances for LAWS. And like the urban planner, that person will know that the intersection of that decision, the environment, the programming, and the actions of other humans will result in a death that would have been preventable had the programmer and engineers been willing to accept the costs. This is not a condemnation of those decisions; perhaps the tradeoff is morally justified. Nor should these claims be mistaken as advocating for a vulgar consequentialism or decision-making that is naively technocratic. Rather, I want to emphasize that these decisions to distribute risk in certain ways—whether in the context of LAWS, self-driving cars, or streetlights—are essentially political and ought to be subject to appropriate principles of justice. However, unless one is willing to say that street lights should be replaced by less reliable, more expensive traffic cops in the name of human dignity, it is hard to see precisely what the objection to LAWS is supposed to be.[24]

Conclusion

My conclusion is a tentative one. Current arguments that purport to show why LAWS are uniquely unjust weapons of war fail. They fail because they compare an idealized human warfighter to the likely messy reality of a LAWS. However, when we compare LAWS to the efforts of militaries that are engaged in sincere, good faith efforts to protect civilians, we have good reason to think that the introduction of autonomous military systems will represent a net improvement. For this reason, researchers are justified in continuing to work on the programming and engineering problems associated with creating sufficiently competent LAWS as to make good that potential. Despite this tentative conclusion, more work needs to be done. There may be good reasons to think that weapons research is problematic in general or that these weapons will inevitably 'fall into the wrong hands' and thus should be prophylactically banned. This paper is silent on those questions.

Acknowledgements I would like to extend my thanks to the editors of this issue, two anonymous reviewers, and the participants of the Moral Technologies Conference.

References

Articles and Books

Alexander, E. P. (1907). *Military memoirs of a confederate*. New York: Skyhorse Publishing.

Angers, T. (2014). *The forgotten hero of My Lai*. New York: Acadian Publishing.

Asaro, P. (2012). On banning autonomous weapon systems: Human rights, automation, and the dehumanization of lethal decision-making. *International Review of the Red Cross, 94*, 687–709.

Barstow, A. (2000). *War's Dirty Little Secret: Rape, Prostitution, and Crimes Against Women*. Pilgrim Press.

Biddle, S. (2006). *Military power: Explaining victory and defeat in warfare in modern battle*. Princeton: Princeton University Press.

Boland, R. (2007). Developing reasoning robots for today and tomorrow. *Signal, 61*, 43.

Brooks, F. Jr. (1987). No silver bullet: Essence and accidents of software engineering. *IEEE Computer, 20*, 10–19.

Chenoweth, E., & Stephan, M. J. (2011). *Why civil resistance works*. New York: Columbia University Press.

Gartner, S., & Segura, G. (1998). War, casualties, and public opinion. *The Journal of Conflict Resolution, 42*(3), 278–300.

[24] Street lights and self-driving cars are not, however, designed in order to take life even if they make choices that take it. Perhaps these human dignity concerns only apply to killings that are done by autonomous systems that are intentionally created to take life. Notice, however, that we are no longer talking about particular actions or decisions made by some autonomous system but are focusing on the telos or purpose of the system. Yet, what constitutes the purpose of a particular technology is notoriously difficult to pin down. We can, of course, describe the streetlight's purpose as 'directing traffic' and a LAWS' purpose as 'killing' but we could also describe the streetlight's purpose as "the provision of transportation-related goods at an acceptable level of traffic-related deaths" and LAWS as "the provision of physical security goods through deterrence at an acceptable level of casualties." This view would also have the perverse implication that commercial autonomous systems that were misused in order to kill would be subject to lesser moral constraints than weapon systems that could be deliberately designed to minimize civilian casualties. At any rate, much would need to be done to make this objection work. The relevant level of description would need to be set, an account of how to precisely determine the function of a weapon system would need to be created, and one would need to show why the function of a system is morally relevant. All of these steps are fraught with difficulty, and this strategy seems to be ultimately contrary to contemporary developments in warfare. In a world where insurgents increasingly use civilian technologies and the line between warfare and law enforcement is increasingly blurred, it seems like we need to develop concepts that do not assume a sharp rupture between military systems and everything else.

Horowitz, M., & Sharre, P. (2015). Meaningful human control in weapon systems: A primer. Center for New American Century Working Paper (Project Ethical Autonomy).
Johnson, A. M., & Axinn, S. (2013). The morality of autonomous robots. *Journal of Military Ethics, 12*(2), 129–141.
Kasher, A., & Yadlin, A. (2005). Assassination and preventive killing. *SAIS Review of International Affairs, 25*, 41–57.
Larson, E. (1996). Casualties and consensus: The historical role of casualties in support of U.S. military operations. Washington, D.C.: RAND Corporation.
Mackinnon, C. (1994). Rape, genocide, and women's rights. *Harvard Women's Law Journal, 17*(5), 5.
Mill, J. S. (2008). *Principles of political economy and chapters on socialism, edited by Jonathan Riley*. Oxford: Oxford University Press.
Pollack, K. (2002). *Arabs at war: Military effectiveness 1948–1991*. Lincoln: University of Nebraska Press.
Rousseva, V. (2004). Rape and Sexual Assault in Chechnya. *Culture, Society, and Praxis*, *3*(1), 64–67.
Schwartz, S. (1994). Rape as a weapon of war in former Yugoslavia. *Hastings Women Law Journal*, *5*(1), 69–74.
Sikkink, K. (2011). *The justice cascade*. New York: WW Norton and Company.
Simpson, T. W., & Mullers, V. C. (2016). Just war and robots' killings. *The Philosophical Quarterly*, *33*, 302–322.
Singer, P. W. (2009). *Wired for War: The Robotics Revolution and Conflict in the 21st Century*. Penguin Books.
Smith, A. M. (2012). Attributability, answerability, and accountability: In defense of a unified account. *Ethics, 122*, 575–589.
Sparrow, R. (2007). Killer robot. *Journal of Applied Philosophy, 24*(1), 62–77.
Swinton, E. D. (1986). *The Defense of Duffer's Drift*. Wayne NJ: Avery Publishing Group.
Walzer, M. (2006). *Just and unjust wars* (4th ed.). New York: Basic Books.
Watson, G. (1996). Two faces of responsibility. *Philosophical Topics, 24*, 227–248.

Reports

Department of the Army. (1998). Field manual 71-1: Tank and mechanized infantry company team.
Human Security Report Project. http://www.hsrgroup.org/human-security-reports/2013/overview.aspx.
International Committee for Red Cross (Autonomous Weapon Systems: Implications of Increasing Autonomy in the Critical Functions of Weapons—ref. 4283-ebook).
Making the case: The dangers of killer robots and the need for a preemptive ban. Human Rights Watch, December 2016.
Office of the Department of Defense. (2005). Development and utilization of robotics and unmanned ground vehicles.