Beyond the Ban: Comparing the Ability of Autonomous Weapon Systems and Human Soldiers to Comply with IHL

Lena Trabucco and Kevin Jon Heller

Much of the legal debate over the use of autonomous weapons systems (AWS) focuses on whether AWS are capable of respecting basic principles of international humanitarian law (IHL). In one camp are the "techno-pessimists:" scholars who insist that AWS are not and never will be capable of complying with IHL. In the other camp are the "techno-optimists:" scholars who believe not only that AWS will eventually be able to comply with IHL, but also that the use of AWS can make armed conflict less violent. There is, however, a puzzling

Kevin Jon Heller *is currently Professor of International Law and Security at the University of Copenhagen's Centre for Military Studies and Professor of Law at the Australian National University. His books include The Nuremberg Military Tribunals and the Origins of International Criminal Law (OUP, 2011) and four co-edited volumes: The Handbook of Comparative Criminal Law (Stanford, 2010), The Hidden Histories of War Crimes Trials (OUP, 2013), The Oxford Handbook of International Criminal Law (OUP, 2018), and Contingency in International Law: On the Possibility of Different Legal Histories (OUP, 2021). He is an Academic Member of Doughty Street Chambers in London and currently serves as Special Adviser to the ICC Prosecutor on International Criminal Law Discourse.*

Lena Trabucco *is a postdoctoral researcher at the University of Copenhagen's Centre for Military Studies and a research affiliate at the Centre for the Study of Existential Risk at Cambridge University. Her research focuses on the strategic and legal implications of emerging military technology, with a particular emphasis on artificial intelligence and quantum technologies. She holds a Ph.D. in political science from Northwestern University and a Ph.D. in law from the University of Copenhagen.*

gap at the heart of this debate. For all their differences, techno-pessimists and techno-optimists agree that the central question is whether AWS will ever be able to comply with IHL better than human soldiers. Yet the debate focuses almost exclusively on the technological limits of AWS; discussion of the cognitive errors that so often cause human soldiers to violate IHL is generally ignored, and systematic comparison of AWS and human soldiers in terms of their ability to comply with IHL is completely absent. Using two scenarios involving the principle of distinction, this short article outlines the kind of research that is needed to determine whether AWS will ever be able to comply with IHL better than human soldiers. It suggests that, given human cognitive limits and machine technological potential, there is indeed reason to believe that in at least in some combat situations AWS will eventually be able to outperform human soldiers in terms of IHL compliance.

INTRODUCTION

In 2013, a coalition of human rights organizations founded the "Campaign to Stop Killer Robots." The Campaign supports a treaty that would ban both "autonomous weapons systems that do not allow for meaningful control" and "autonomous weapons systems that would target humans, even when used with human control."[1] Nearly seventy states have indicated their support for such a treaty,[2] and the Campaign has had notable success in increasing public opposition to AWS: recent polling indicates that more than 60 percent of individuals in twenty-six countries categorically oppose their use.[3]

The Campaign to Stop Killer Robots has failed, however, to convince the states that have been most active in developing AWS: Israel, Russia, China, the United States, and the United Kingdom.[4] Not only have those states refused to support a treaty banning the use of AWS, they also continue to invest heavily in them. China is pouring considerable resources into the technology behind drone swarms, a network entity using sophisticated artificial intelligence (AI) to detect human body heat.[5] The United States spent USD 18 billion on AWS research between 2016 and 2020.[6] And Russia has declared its intention to have one-third of its military run by AI no later than 2030.[7]

While systems like drone swarms seem to belong more to science fiction than the modern battlefield, there is considerable evidence that offensive AWS[8] will be regularly used in armed conflict—sooner than we think. Reports indicate, for example, that both Russia and Ukraine have deployed weapons with fully autonomous functions in their current international

armed conflict. Russia has used Lancet drones that, once launched, circle a predetermined geographic area and then engage a preselected target without human intervention.[9] Ukraine has successfully used Punisher drones, which possess autonomous targeting capabilities when used in tandem with a smaller reconnaissance drone, to target stationary Russian facilities such as fuel storages, ammunition reserves, and cyberwarfare centers.[10] It is unclear whether Lancets or Punishers have been used fully autonomously, because both types of drones are normally used with a human supervising the targeting missions, but such use cannot be ruled out.[11]

The Campaign to Stop Killer Robots' primary objection to AWS is ethical—that allowing machines to use lethal force is incompatible with human dignity.[12] The most common argument for banning AWS, however, is legal: that autonomous weapons are incapable of respecting the basic principles of IHL. The International Committee of the Red Cross (ICRC),[13] for example, argues that "the use of AWS to target human beings should be ruled out" because "[e]ffectively protecting combatants/fighters who are placed *hors de combat* and civilians who are not, or no longer, taking a direct part in hostilities calls for difficult and highly contextual, conduct-, intent- and causality-related legal assessments by humans in the context of a specific attack."[14]

As the quote from the ICRC indicates, the legal argument against AWS is based on two assumptions: one explicit and one implicit. The explicit assumption is that AWS are not only currently incapable of complying with IHL—they will never be able to comply. The implicit assumption is that, unlike AWS, human soldiers can reliably comply with IHL.

Neither assumption is self-evident. The explicit assumption concerning the limits of AWS may be correct for currently existing systems, but not even the most cutting-edge roboticists are willing to confidently predict how AWS technology might evolve. And the implicit assumption about the superiority of human judgment is difficult to reconcile with statistics that indicate civilian death and destruction are endemic to modern armed conflict. Nearly half of all civilian casualties during the war in Afghanistan, for example, were caused by human misidentification.[15]

To be sure, these considerations do not mean the ICRC is wrong. It is possible AWS will never be technologically sophisticated enough to respect basic principles of IHL, such as distinction. It is also possible that even if states do develop AWS with that ability, human soldiers will still outperform them. But those are empirical questions, not ethical or legal ones, and they can be answered only by systematically comparing the ability of AWS and human soldiers to comply with IHL.

Recent academic discourse on AWS has barely scratched the surface of this comparison. Discussion of the ability of AWS to comply with IHL has focused almost exclusively on the current state of technology, generally ignoring how the astronomical pace of AI innovation and sensor development is likely to dramatically improve compliance. Discussion of the ability of human soldiers to comply with IHL has generally ignored what cognitive psychology has to say about decision-making in situations of risk and uncertainty. And systematic comparison of the two in terms of IHL compliance is entirely non-existent.

This short article cannot begin to rectify the deficiencies in debates over the legal regulation of AWS. Its purpose is far more modest: to provide an overview of the kind of research that is needed to determine whether AWS can ever be expected to comply with IHL better than human soldiers. Section 1 begins by mapping the AWS debate through the lens of "techno-optimism" and "techno-pessimism." Section 2, the heart of the article, then uses two scenarios involving the principle of distinction to outline what a compelling comparison of the ability of AWS and humans to comply with IHL would look like. Finally, Section 3 offers some thoughts about why the research agenda we call for in the article is so important.

SECTION 1: THE TECHNO-PESSIMISTS AND TECHNO-OPTIMISTS

Scholars are deeply divided over the need for an international treaty to ban AWS. Those in the "techno-pessimist" camp argue that banning AWS is both ethically and legally necessary.[16] The ethical argument is rooted in the claim that only humans should be responsible for taking human life in armed conflict.[17] This ethical concern is particularly important to the techno-pessimist position because many scholars in that camp believe it provides a stronger argument for the ban of AWS than IHL compliance. The legal objection to AWS will be irrelevant if technology advances to the point that AWS can respect IHL's basic requirements.[18]

An adequate critique of the ethical argument against AWS is beyond the scope of this article. It is nevertheless worth noting that the argument necessarily assumes AWS will never be able to outperform humans in terms of IHL compliance. If technology ever advances to the point where AWS can comply with IHL better than human soldiers, banning their use would reflect a deliberate decision to prefer more death to less death simply because humans are doing the killing instead of machines. Such a preference would hardly be "ethical."

The legal argument for banning AWS, as noted above, is that AWS

are unable to comply with basic principles of IHL.[19] For example, Noel Sharkey—probably the most-cited roboticist on the IHL issue—has claimed that "autonomous armed robot systems cannot discriminate between combatants and non-combatants or other immune actors such as service workers, retirees, and combatants that are wounded, have surrendered, or are mentally ill in a way that would satisfy the principle of distinction."[20] He offers three reasons for his techno-pessimism. First, AWS do not have the "adequate sensory or vision processing systems" necessary to identify whether an individual is a targetable combatant or if an individual who falls into one of the category protected against attack, because the systems "may be able to tell us that something is a human, but they could not tell us much else."[21] Second, IHL lacks an "adequate definition of a civilian that we can translate into machine code."[22] Third, even if sensory systems improve over time, providing AWS with the "battlefield awareness or common sense reasoning" that is required by the principle of discrimination is likely "computationally intractable."[23]

Scholars in the "techno-optimist" camp tell a different story. These scholars see no reason why states cannot develop AWS that can comply with IHL. Indeed, some techno-optimists, most notably the roboticist Ron Arkin, go even further. In Arkin's view, given human soldiers' "dismal record" concerning IHL compliance,[24] the development of advanced AWS could "result in a reduction in collateral damage and damage to civilian property, which translates into saving innocent lives," making their use "a moral requirement."[25]

Arkin provides six reasons for his optimism. First, AWS do not have an instinct for self-preservation, which means they can use a "first do no harm" approach to targeting, acquiring as much information as possible about targets, even if it means the AWS will be destroyed. Frightened soldiers, by contrast, are likely to take a far more dangerous "shoot first, ask questions later" approach.[26] Second, AWS will eventually possess sensing capabilities that are far superior to human vision, making them much more useful for piercing the "fog of war."[27] Third, AWS are not subject to the emotional states—hate, fear, hunger, exhaustion, stress—that so often cause human soldiers to violate IHL.[28] Fourth, AWS are not subject to "scenario fulfillment," a cognitive bias that leads human soldiers to bring preconceived expectations about an adversary into combat and then disregard new information or data that is inconsistent with those expectations.[29] Fifth, AWS can process and integrate information more quickly and from a far greater range of sources than humans, which is essential to making accurate targeting decisions given the fluidity and unpredictability of modern

armed conflict.[30] Sixth, and finally, when integrated into a machine/human team, AWS can serve as objective and incorruptible reporters of IHL violations committed by human soldiers.[31]

SECTION 2: COMPARING HUMANS AND MACHINES

Techno-pessimists reject many of the arguments in favor of AWS. The central point, however, is this: although we cannot be certain what AWS will be capable of, even in the near future, it is at least *possible* that AWS will eventually be able to comply with IHL better than human soldiers. Given that possibility, efforts to "stop killer robots" are not only impractical—because powerful states continue to pour funds into AWS development and are thus exceedingly unlikely to join a treaty prohibiting their use[32]—they are also unethical, because they might ensure, however unintentionally, that modern armed conflict becomes more violent than necessary.

It is time, in short, to move beyond the ban. Instead, we need ambitious empirical research to identify the kinds of combat situations in which AWS can be expected—either now or in the future—to comply with IHL better than human soldiers. Such research would involve six steps for each relevant principle of IHL:

1. Identifying the specific cognitive tasks required for a combatant—machine or human—to reliably comply with the principle.
2. Identifying the current ability of AWS to complete those cognitive tasks.
3. Projecting how AWS ability might improve over time.
4. Identifying the current ability of human soldiers to complete those cognitive tasks.
5. Determining whether human cognitive performance can be improved over time.
6. Mapping, based on the results of steps two through five, the situations in which AWS can be expected to comply with the principle better than human soldiers—and vice-versa.

To get a sense of what this kind of research would entail, consider two scenarios involving the principle of distinction, the requirement that parties to an armed conflict shall "at all times distinguish between the civilian population and combatants and between civilian objects and military objectives and accordingly shall direct their operations only against military objectives."[33] The first scenario is relatively simple: an attacker spots a civilian truck loaded with building supplies that is parked in an area where enemy forces are operating. An object is a military objective

if it contributes to the enemy's military action through its nature, location, purpose, or use.[34] The truck is not a military objective because none of those criteria are satisfied, so it cannot be lawfully targeted. To avoid unlawfully targeting the truck, therefore, a combatant (human or machine) must successfully complete the following cognitive tasks:

1. Identify the object as a vehicle, as opposed to, say, a gun emplacement (a military objective by nature).
2. Identify the vehicle as a truck, as opposed to, say, a tank (a military objective by nature).
3. Identify the truck as civilian, not military (a military objective by nature).
4. Identify the civilian truck as loaded with building supplies for civilian infrastructure instead of with, say, ammunition (a military objective by use).

The second scenario is more complex: an attacker spots a group of civilian women using metal sickles to cut grass in an area where enemy forces had previously carried out attacks. The women cannot lawfully be targeted because they are not combatants—neither members of the enemy's regular armed forces nor members of irregular armed forces—and are not civilians directly participating in hostilities. To avoid unlawfully targeting the women a human or machine combatant must successfully complete the following cognitive tasks:

1. Identify the women's clothing as ordinary civilian garb, not as a uniform (indicating membership in regular armed forces).
2. Avoid identifying anything the women are wearing as a fixed and distinctive sign (indicating membership in irregular armed forces).
3. Identify the women as cutting grass, as opposed to, say, building entrenched positions (direct participation in hostilities).
4. Identify the objects the women are carrying as metal sickles instead of, say, rifles (suggestive of direct participation in hostilities).

In the absence of the necessary research, it is not possible to conclusively determine which of the two, human soldiers or AWS, would be more likely to comply with the principle of distinction in the two scenarios. What we know about the limits of human cognition, however, suggests that AWS might outperform human soldiers—at least eventually.

A. Human Soldiers

Statistics concerning fratricide—incidents in which soldiers target their own side instead of the enemy—indicate that, in general, human

soldiers are prone to make serious identification errors in the heat of battle. During Operation Desert Storm, for example, friendly fire accounted for 24 percent of American casualties[35] and 77 percent of the total number of U.S. armored vehicles destroyed—27 of 35 Bradley fighting vehicles and M1 Abrams tanks.[36]

Research indicates that 80-85 percent of these kinds of military accidents are caused by errors of human cognition.[37] That is not terribly surprising given what we know about how dangerous and stressful situations like combat affect human decision making. Consider six factors that are endemic to most, if not all, armed conflict: noise, heat, time pressure, sleep deprivation, performance pressure, and stress. Noise significantly degrades performance of combat tasks, such as disposing of explosive ordinance,[38] while heat negatively affects soldiers' ability to estimate time and read maps.[39] Time pressure has a similar effect on soldiers as noise because of "the cognitive demands, or information overload, imposed by the requirement to process a given amount of information in a limited amount of time."[40] The effects of sleep loss are even greater, degrading cognitive performance in general as well as visual vigilance, situational awareness, reaction time, and mood.[41] Performance pressure on soldiers, whether caused by the need to avoid making mistakes that might lead to fellow soldiers being harmed or incurring the wrath of commanders, "tends to impair performance on difficult tasks."[42] And finally, stress—particularly stress caused by the threat of harm—has the most damaging effects of all, not only leading to "perceptual tunneling, reduced working memory, and performance rigidity," but also making soldiers prone to what is known as "ballistic decision making," i.e., "making decisions without checking the consequences."[43]

This cognitive-psychological research indicates that human soldiers could easily make identification errors in the two distinction scenarios presented above. The loss of concentration caused by the stress of combat, excessive heat, or lack of sleep, for example, could cause soldiers to perceive the civilian truck as a military one. As Keebler notes, because a soldier "must focus most or all of his or her attention to properly make target identification decisions," if the soldier "cannot provide adequate attention to a potential threat vehicle to allow for a temporary object representation to form, then the soldier will be unable to make correct judgments as to the identity and alliance of the given vehicle,"[44] with potentially fatal consequences.

Similarly, human soldiers could also misidentify the truck because of cognitive overload. Soldiers are almost always required to perform more

than one task at a time in combat; a tank gunner, for example, "must simultaneously identify targets, operate their gun, and listen to the tank commander's instructions."[45] The problem with such "dual tasking" is that it significantly reduces the accuracy of target identification, because it requires soldiers to divide their limited attentional capacity.[46] Indeed, human ability to focus on more than one task at a time is so limited that "[a]ny other task involving even minimal amounts of attention could diminish perceptual resources needed to make correct identifications."[47]

Finally, misidentification could occur because of what has been called the "friend/foe bias." Research indicates that, during combat, human soldiers are more likely to identify an unknown vehicle as belonging to the enemy than to their own armed forces. "In a tactical situation," Briggs notes, "[t]he battlefield gunner faces a great deal of uncertainty and expects to observe the enemy."[48] That expectation, in turn, makes it more likely that the gunner will focus on aspects of the unknown vehicle that confirm her perception that she is seeing an enemy vehicle.[49]

Many, if not all, of the same concerns apply equally to the second distinction scenario. Stress-related loss of concentration or cognitive overload caused by dual tasking could lead human soldiers to misperceive the women's civilian clothing as military uniforms, the grass cutting as building an entrenched position, or holding sickles as holding weapons. Similarly, the friend/foe bias could cause soldiers to wrongly assume that the women were combatants or civilians directly participating in hostilities, particularly given that their compatriots had been previously attacked in the area.

Another type of cognitive error, "object use bias," could also lead human soldiers to make distinction errors in the farming scenario. A group of researchers hypothesized that an individual's perception of objects being held by others could be affected by the nature of the object she herself is carrying. The results of their experiments were unequivocal:

> "The familiar saying goes that when you hold a hammer, everything looks like a nail. The apparent harmlessness of this expression fades when one considers what happens when a person holds a gun. We have shown here that, having the opportunity to use a gun, a perceiver is more likely to classify objects held by others as guns and, as a result, to engage in threat-induced behavior (in this case, raising a firearm to shoot)."[50]

The relevance of this research to the farming scenario is apparent. Because human soldiers carry weapons, object-use bias indicates that they will be more likely to perceive objects held by potential enemy combatants as weapons than they would if they were unarmed.

B. AWS

There are, in short, a variety of reasons to question the ability of human soldiers to reliably comply with the principle of distinction in the two scenarios described above. In fact, American soldiers did violate the principle in the farming scenario, which occurred during the conflict in Afghanistan. The soldiers made notable cognitive errors: inferring hostile intent from previous attacks on Americans in the area; assuming the farmers were men from the absence of headscarves (which the women were using to transport the cut grass); and misperceiving the sickles as guns. As a result, wrongly concluding they were about to be attacked, the soldiers killed one civilian woman and injured four, all of whom were between the ages of six and seventeen.[51]

Would AWS be more likely to reliably comply with the principle of distinction? In one respect, they are clearly superior: as noted earlier, unlike human soldiers, AWS are impervious to noise and heat, don't need to sleep, feel no pressure from time or their compatriots, and do not experience the stress of knowing that they could be killed at any moment. Given that those factors are each likely to play a significant role in distinction-related mistakes, AWS would have some inherent advantages over human soldiers in *any* situation requiring the ability to distinguish between military objectives and civilian objects or between combatants, civilians directly participating in hostilities, and civilians.

There are also more specific reasons to believe that AWS would eventually be able to outperform human soldiers in the two scenarios. The reliability of human identification of civilian vehicles and civilians is likely to be negatively affected by both dual tasking and friend/foe bias. AWS do not experience cognitive overload no matter how many tasks they are called upon to perform simultaneously. They also will not assume that a vehicle belongs to the military or that civilians are directly participating in hostilities simply because they expect to find the enemy in a combat area. Moreover, an AWS would not assume civilians are armed simply because they are armed themselves.

That said, complying with the principle of distinction requires AWS to do more than avoid certain cognitive errors that tend to afflict human soldiers. As Winter has pointed out, "machines would [also] need to possess advanced skills in observation and recognition as well as sophisticated judgment-making ability."[52]

The first skill, observation, is clearly a strength of AWS. Not only can AWS continue to observe a combat situation much longer than humans

because they neither tire nor need to preserve themselves, but machine vision is also vastly more sensitive than human vision. The hybrid automated/autonomous Tomahawk missile, for example, uses "digitized scene mapping" to target static objects with accuracy that "far exceeds the capability of humans."[53] Similarly, Raytheon and Exyn Technologies have developed "a fully autonomous aerial robot" that "can operate in GPS-denied environments to map dense urban environments in 3-D" and can "dig deep to reveal tunnels, urban undergrounds and natural cave networks."[54] Indeed, the robot's vision system, which relies on a combination of cameras, radar, and lidar (or light detection and ranging), is so sensitive that it can detect "even dangling wires."[55]

Current AWS would have trouble with the second skill, object recognition, in many combat situations—particularly those that take place in dense urban environments. Object-recognition algorithms use computer vision systems, and those systems "struggle…to integrate incoming sensor data into the 'brain' of an autonomous system to form a world model by which action decisions can be made."[56] That said, object-recognition software built for dynamic environments like a battlefield has improved dramatically over the past few years,[57] to the point where AWS could already significantly outperform human soldiers in certain combat functions. For example, Patriot One's "PatScan" security-guard system, which is designed to automatically detect when a person passing through a doorway or turnstile is carrying a weapon, is able to detect both overt and concealed weapons and can even identify the type of weapon a person is carrying with nearly 95 percent accuracy.[58] Although PatScan is not currently used in the AWS context, the military use of the system is clear.

The third and final skill, judgment, is perhaps the most challenging for AWS. As Winter notes, "in spite of the successes in machine observation and recognition, machine judgment still appears to be lagging behind that of humans."[59] That lag can indeed limit the ability of AWS to comply with the principle of distinction. Even techno-optimist scholars like Sassoli, for example, acknowledge that it "may be particularly difficult to automate the indicators that convince a human being a certain person belongs to a category (combatants; possibly members of an organized armed group who assume a continuous combat function) or is engaged in conduct (direct participation in hostilities) which makes them a legitimate target."[60]

Even here, though, qualifications are necessary. The first is that scholars often overstate the range of combat situations in which AWS will likely never be able to comply with the principle of distinction. Thus, Umbrello and Wood wrongly claim that AWS cannot be used on the battlefield

because they cannot reliably determine when a combatant has become *hors de combat* through capture, incapacitation, or surrender.[61] Contrary to their argument, determining whether a combatant is attempting to surrender, for example, does not require sensitivity to "contextual factors."[62] It simply requires recognizing what kinds of actions "clearly express an intention to surrender" for purposes of IHL, such as, raising a white flag or throwing hands in the air.[63] Such recognition will eventually be well within the capabilities of AWS, if it isn't already.

The second qualification is that there is a straightforward fix for situations in which AWS cannot comply with the principle of distinction: "allowing a weapon system to target autonomously only those categories… that are, without question, targetable."[64] That fix might be unsatisfying in terms of AWS's effectiveness, but it would essentially guarantee that AWS will not violate the principle of distinction.

The third qualification, which cannot be stressed enough, is that human soldiers are not necessarily any better at using the kind of judgment the principle of distinction requires. Krishnan, for example, claims that AWS may never be able to reliably distinguish between combatants and civilians because "it would be easy for terrorists or insurgents to trick these robots by concealing weapons."[65] Schmitt and Thurnher's response is on point: "asymmetrically disadvantaged enemies have been feigning civilian or other protected status to avoid being engaged by human-operated weapon systems for centuries."[66] To be sure, it is possible that human soldiers are now and always will be better than AWS at detecting such perfidy. But that is once again an empirical question that can only be answered through the kind of comparative research we are calling for in this article.

The fourth and final qualification is related to the third: although AWS technology will continue to improve over time, reducing and perhaps eventually even eliminating the kinds of distinction problems techno-pessimists cite, the cognitive limitations of human soldiers will persist. Certain methods exist to debias human judgment, but they all require training individuals to engage in reasoned analysis, instead of relying on their intuitions.[67] Those methods, therefore, are particularly ill-suited to improving the judgment of human soldiers in combat situations: as the U.S. Army points out in a white paper on cognitive psychology, "[t]he complexity and uncertainty of these situations may not afford individuals the luxury of time or collaboration that the methods require, particularly for individuals operating at the tactical level where commanders and soldiers must rely on quick, often inherently intuitive, decisions."[68]

CONCLUSION: THE IMPORTANCE OF THE RESEARCH

There is a puzzling gap at the heart of legal debate over whether autonomous weapons systems can comply with international humanitarian law. For all their differences, techno-pessimists and techno-optimists agree that the relevant comparison is between AWS and human soldiers. Yet the debate between them focuses almost exclusively on the technological limits of AWS; discussion of the cognitive limits of human soldiers is minimal, and systematic comparison of AWS and human soldiers is completely absent.

As we have demonstrated, this gap is problematic. Despite the efforts of well-meaning movements like the Campaign to Stop Killer Robots, the machine has left the hangar: states will continue to develop AWS and will deploy them in combat as soon as their military advantage becomes clear. It is thus imperative for states to have a precise understanding of the kinds of combat situations in which AWS are likely to comply with IHL better than human soldiers, and vice versa. Such understanding is particularly essential at the tactical level: a military concerned about protecting civilians and civilian objects from unnecessary harm will always want to choose the means of attack—AWS or human soldiers—most likely to comply with IHL. Indeed, given that an attacker must take "all feasible precautions in the choice of means and methods of attack with a view to avoiding, and in any event to minimizing, incidental loss of civilian life, injury to civilians and damage to civilian objects,"[69] states that possess AWS will be legally required to do so.

There is, however, a less apparent but equally important reason to pursue the kind of comparative research we have called for in this article: it will help states decide how to direct their research into AWS. Simply put, law-abiding states have no incentive to commit precious resources to developing types of autonomous weapons that will always be inferior to human soldiers in terms of IHL compliance. Their research dollars are far better spent on AWS that have a reasonable chance of outperforming human soldiers.

This article has focused on only one of many basic principles of IHL, distinction, and it has provided merely an impressionistic analysis of the ability of AWS and human soldiers to comply with that principle. Nevertheless, given the significant cognitive limits human soldiers face, even such an impressionistic analysis provides reason to believe that AWS will eventually be able to comply with IHL better than human soldiers, at least in some combat situations. If we care about minimizing unnecessary violence in armed conflict, that possibility alone justifies moving beyond the ban. ***f***

ENDNOTES

1 "Our Policy Position," *Stop Killer Robots*, https://www.stopkillerrobots.org/our-policies/ (accessed April 20, 2022).
2 "A critical opportunity to ban killer robots – while we still can," *Amnesty International*, November, 2021, https://www.amnesty.org/en/latest/news/2021/11/global-a-critical-opportunity-to-ban-killer-robots-while-we-still-can/.
3 "Global Survey Highlights Continued to Fully Autonomous Weapons," *Ipsos,* February 2, 2021, https://www.ipsos.com/en-us/global-survey-highlights-continued-opposition-fully-autonomous-weapons.
4 Noel E. Sharkey, "The evitability of autonomous robot warfare," *International Review of the Red Cross* 94 (886) (Summer 2012): 788.
5 Ibid.
6 James Dawes, "What you need to consider about 'killer robots' and autonomous weapons research," *Fast Company*, December 29, 2021, https://www.fastcompany.com/90707966/what-you-need-to-consider-about-killer-robots-and-autonomous-weapons-research.
7 David Freedman, "US Is Only Nation with Ethical Standards for AI Weapons. Should We Be Afraid?" *Newsweek*, September 15, 2021. https://www.newsweek.com/2021/09/24/us-only-nation-ethical-standards-ai-weapons-should-we-afraid-1628986.html.
8 Autonomous weapons, such as the Phalanx close-in weapons system, have long been used for defensive purposes. *See* generally Vincent Boulanin and Maaike Verbruggen, *Mapping the Development of Autonomy in Weapons Systems* (Solna: SIPRI, 2017), https://www.sipri.org/sites/default/files/2017-11/siprireport_mapping_the_development_of_autonomy_in_weapon_systems_1117_1.pdf.
9 Jeremy Kahn, "A.I. is on the front lines of the war in Ukraine," *Fortune*, March 1, 2022, https://fortune.com/2022/03/01/russia-ukraine-invasion-war-a-i-artificial-intelligence/.
10 Haye Kesteloo, "Punisher Drones are Positively Game-Changing for Ukrainian Military in Fight Against Russia," *DroneXL*, March 3, 2022, https://dronexl.co/2022/03/03/punisher-drones-ukrainian-military/.
11 Kahn, "A.I. is on the front lines"; Kesteloo, "Punisher Drones." There are reports that similar technology was used in the 2020 conflict between Azerbaijan and Armenia. *See* Kelsey Atherton, "Loitering munitions preview the autonomous future of warfare," *Tech Stream* (blog), *Brookings Institution,* August 4, 2021, https://www.brookings.edu/techstream/loitering-munitions-preview-the-autonomous-future-of-warfare/.
12 Amnesty International, "A critical opportunity."
13 The ICRC is "an independent, neutral organization ensuring humanitarian protection and assistance for victims of armed conflict and other situations of violence." ICRC, "Mandate and Mission," https://www.icrc.org/en/who-we-are/mandate.
14 ICRC, "International Committee of the Red Cross (ICRC) position on autonomous weapon systems: ICRC position and background paper," *International Review of the Red Cross* 102 (915) (2020): 1345.
15 Larry Lewis, *Redefining Human Control: Lessons from the Battlefield for Autonomous Weapons* (Arlington: CAN, 2018), 4, https://www.cna.org/cna_files/pdf/DOP-2018-U-017258-Final.pdf.
16 Campaign to Stop Killer Robots, https://www.stopkillerrobots.org/.
17 *See* Elvira Rosert and Frank Sauer, "Prohibiting Autonomous Weapons: Put Human Dignity First," *Global Policy* 10 (3) (September 2019): 370.
18 Ibid.

19 *E.g.,* Noel Sharkey, "Saying 'No!' to Lethal Autonomous Targeting," *Journal of Military Ethics* 9 (4) (2010): 378. This position has also been endorsed by some states. *E.g.*, Austria, Brazil, Chile, Ireland, Luxembourg, Mexico, and New Zealand submitted a collective statement to the Group of Governmental Experts (GGE) that argued, inter alia, that "[t]he application of and compliance with IHL requires context-specific, value-based judgment by human beings and this requirement cannot be substituted by autonomous machines or systems." *See* Group of Governmental Experts on Lethal Autonomous Weapons Systems, "Joint Submission on possible consensus recommendations in relation to the clarification, consideration and development of aspects of the normative and operational framework on emerging technologies in the area of lethal autonomous weapons systems, submitted by Austria, Brazil, Chile, Ireland, Luxembourg, Mexico, and New Zealand," (June 2021): 1, http://www.converge.org.nz/pma/jsub-gge,june21.pdf.
20 Sharkey, "Evitability," 788.
21 Ibid.
22 Ibid., 789.
23 Ibid.
24 Ronald Arkin, "Lethal Autonomous Systems and the Plight of the Non-combatant," *AISB Quarterly* 137 (July 2013): 2.
25 Ibid.
26 Ibid., 3.
27 Ibid.
28 Ibid.
29 Ibid.
30 Ibid., 3-4.
31 Ibid., 4.
32 Proponents of a ban on AWS often cite the success of the Ottawa Convention, which prohibits the use of anti-personnel landmines and has more than 160 states parties, as precedent for a similar treaty regarding AWS. *See*, e.g., Rosert and Sauer, "Dignity," 371. The Convention's success, however, is something of a pyrrhic victory: nearly every state that has produced, traded, or used anti-personnel landmines have either refused to ratify the Convention or rejected calls for a categorical ban. Richard Price, "Emerging Customary Norms and Anti-Personnel Landmines," in *The Politics of International Law*, ed. Christian Reus-Smit (Cambridge: Cambridge University Press, 2009), 121. There is no reason to believe a treaty banning AWS will meet a better fate.
33 Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of International Armed Conflicts (Protocol I), June 8, 1977, 1125 U.N.T.S. 3, entered into force December 7, 1978, Art. 48.
34 Protocol I, Art. 52(2).
35 Joseph R. Keebler et al., "A Cognitive Basis for Friend/Foe Misidentification of Vehicles in Combat," in Dee H. Andrews, Robert P. Herz, and Mark B. Wolf, eds., *Human Factors Issues in Identification* (Surrey: Ashgate, 2010), 131.
36 Ralph W. Briggs, "Battlefield recognition of armored vehicles," *Human Factors* 37 (3) (September 1995): 596.
37 Lobna Cherif et al., "Multitasking in the military: Cognitive consequences and potential solutions," *Applied Cognitive Psychology* 32 (2018): 430.
38 James E. Driskell, Eduardo Salas, and Joan H. Johnston, "Decision Making and Performance Under Stress," in Thomas W. Britt, Carl Andrew Castro, and Amy B. Adler, eds., *Military Life: The Psychology of Serving in Peace and Combat*, (Westport: Praeger Security International, 2006), 134.

39 Eric R. Muth et al., "Augmented Cognition: Aiding the Soldier in High and Low Workload Environments Through Closed-Loop Human-Machine Interactions," in Britt, Castro, and Adler, *Military Life*, 110.
40 Driskell, Salas, and Johnston, "Decision Making," 135.
41 Muth et al., "Augmented Cognition," 110.
42 Driskell, Salas, and Johnston, "Decision Making," 138.
43 Ibid., 140.
44 Keebler et al., "Friend/Foe," 134.
45 Cherif et al., "Multitasking," 430.
46 Related phenomena affecting a soldier's ability to distinguish combatants from civilians include "task switching," in which a soldier has to frequently change tasks, and "task interruption," in which a soldier must stop engaging in a task to, say, receive instructions from his commander. Both phenomena equally degrade cognitive performance. *See* Ibid.
47 Keebler et al., "Friend/Foe," 134.
48 Briggs, "Battlefield Recognition."
49 Ibid.
50 Jessica K. Witt and James R. Brockmole, "Action Alters Object Identification: Wielding a Gun Increases the Bias to See Guns," *Journal of Experimental Psychology: Human Perception and Performance* 38 (5) (2012): 1165.
51 Lewis, "Redefining Human Control," 5.
52 Elliot Winter, "The Compatibility of Autonomous Weapons with the Principles of International Humanitarian Law," *Journal of Conflict & Security Law* 27 (1) (2022): 13.
53 M.L. Cummings, "Lethal Autonomous Weapons: Meaningful Human Control or Meaningful Human Certification?" *Technology and Society*, December 23, 2019. https://technologyandsociety.org/lethal-autonomous-weapons-meaningful-human-control-or-meaningful-human-certification/.
54 Elliot Winter, "The Compatibility of Autonomous Weapons with the Principle of Distinction in the Law of Armed Conflict," *ICLQ* 69 (October 2020): 858.
55 Ibid. "Lidar" stands for light detection and ranging.
56 Cummings, "Meaningful Human Control."
57 *See* generally Wu Xiongwei, Steven C. H. Hoi, and Doyen Sahoo, "Recent Advances in Deep Learning for Object Detection," *Neurocomputing* 396 (January 2020). While the article does not discuss battlefield settings specifically, it surveys recent improvements in object detection and details the factors that affect detection performance.
58 Winter, "Principles of International Humanitarian Law," 13-14.
59 Ibid., 14.
60 Marco Sassoli, "Autonomous Weapons and International Humanitarian Law: Advantages, Open Technical Questions, and Legal Issues to Be Clarified," *International Law Studies* 90 (2014): 327.
61 Steven Umbrello and Nathan Gabriel Wood, "Autonomous Weapons Systems and the Contextual Nature of *Hors de Combat* Status," *Information* 12 (2021): 9. *See* also Winter, "Principles of International Humanitarian Law," 14.
62 Umbrello and Wood, "Autonomous Weapons Systems," 2.
63 Sassoli, "Autonomous Weapons," 315.
64 Ibid., 327.
65 Armin Krishnan, *Killer Robots: Legality and Ethicality of Autonomous Weapons* (Surrey: Ashgate, 2009), 99.

66 Michael N. Schmitt and Jeffrey S. Thurnher, "'Out of the Loop': Autonomous Weapon Systems and the Law of Armed Conflict," *Harvard National Security Law Journal* 4 (2013): 247.

67 Joseph Rodman, "Cognitive Biases and Decision Making: A Literature Review and Discussion of Implications for the US Army" (January 2015), 22, https://www.fededtv.com/tvwwimages/fedexec/HDCDTF_WhitePaper_Cognitive%20Biases%20and%20Decision%20%20Making_Final_2015_01_09_0.pdf.

68 Ibid.

69 Protocol I, Art. 57(2)(ii).