Lethal Autonomous Weapons Systems:
Can Targeting Occur Without Ethical Decision-Making?

LCDR Andrew M. Coffin, JAGC, USN

United States Naval War College
College of Naval Command and Staff, June 2019
Written in Completion of the Graduate Certificate in Ethics and Emerging Military Technology

Table of Contents

Abstract

The rise in capability of lethal autonomous weapons systems has led to a backlash against their use on both legal and ethical grounds. While the academic literature evinces a general ethical objection to the employment of lethal autonomous weapons, the specific nature of that objection is not clearly explored. Opponents agree that the use of lethal autonomous weapons crosses a moral line; however, they do not elucidate where that line is drawn. This paper argues that the moral line is crossed when lethal autonomous weapons systems are given ethical agency to make life or death decisions free of human input. Furthermore, the legal objections raised to lethal autonomous weapons systems are not inherently legal; instead, they are based on this ethical objection applied to the legal context. Given this underlying objection to lethal autonomous weapons systems being given moral agency, this paper analyzes whether non-ethical lethal autonomous weapons systems—i.e. systems that are not allowed to make any ethical decisions—can comply with international humanitarian law in a way that is still militarily useful. It concludes that non-ethical lethal autonomous weapons systems would be able to comply with international humanitarian law in specific situations, specifically, in near-peer—particularly near-peer maritime—combat.

I. Introduction

Automated weapons systems have been in use by the United States military for decades. As far back as World War II, in Operation Aphrodite, the United States Army Air Forces, in collaboration with the United States Navy, attempted to use remote controlled bomber airplanes filled with explosives against Nazi targets that were too difficult to bomb with manned aircraft.[1] While this particular operation was unsuccessful, it was prophetic of the future of warfare. Over the last two decades, automated weapons systems proliferated and are become increasingly autonomous.

With the abundance of increasingly autonomous weapons systems (AWS) there has been a significant backlash against autonomous lethality.[2] Much of this anti-autonomous lethality literature is couched in legal analysis of whether or not lethal AWS are capable of complying with international humanitarian law (IHL).[3] Rather than a purely legal argument of compliance, however, the legal arguments are actually premised on underlying ethical objections. The use of purely legalistic language obfuscates the ethical objections upon which criticisms of lethal AWS rest.

This paper will show that lethal AWS cross a moral line when made to demonstrate ethical agency. Because giving lethal AWS ethical agency crosses a moral line, they should only be allowed to make non-ethical decisions. If, however, lethal AWS are limited to only non-ethical decision-making, can they still comply with international

---

[1] P. W. Singer, *Wired for War: The Robotics Revolution and Conflict in the Twenty-first Century* (New York: Penguin Books, 2009) 48-49.
[2] Human Rights Watch and International Human Rights Clinic at Harvard Law School, *Losing Humanity: The Case Against Killer Robots* (Human Rights Watch, November 2012) https://www.hrw.org/sites/default/files/reports/arms1112ForUpload_0_0.pdf.
[3] Human Rights Watch, *Losing Humanity*, 1; Ronald Arkin, *Governing Lethal Behavior in Autonomous Robots* (Boca Raton, FL: Chapman and Hall, 2009) 37-48 (discussing other thinkers' objections to lethal AWS).

humanitarian law (IHL)? To answer this question, this paper will parse through the targeting cycle of lethal AWS to identify which aspects of the targeting cycle require an ethical decision—and, therefore, must have a human-in-the-loop to make the decision—and which aspects of the targeting cycle do not require an ethical decision—and, therefore, can be made autonomously. While non-ethical lethal AWS cannot conduct the full range of targeting allowable under IHL, it can both comply with IHL and still conduct some autonomous targeting in militarily relevant ways.

II. Background

Before delving into analysis, it is necessary to clarify terms. The vocabulary surrounding AWS has not yet coalesced into a single lexicon shared by all in the field. For purposes of this paper, the following definitions will be used. First, this paper deals with both autonomous and automated systems. Autonomous systems operate in "dynamic, unstructured, open environments based on feedback information from a variety of sensors."[4] Automated systems, on the other hand, are "unsupervised systems or processes that involve repetitive, structured, routine operations without much feedback information."[5] In other words, autonomous systems are capable of making independent decisions based on the input of their sensors, while automated systems are programed to perform a single function or single series of functions based on a single triggering action. As such, this paper will use Dr. Peter Asaro's comprehensive definition of a lethal autonomous weapon system as "any system that is capable of targeting and initiating the use of potentially lethal force without direct human supervision and direct human

[4] Peter Asaro, "On Banning Autonomous Weapon Systems: Human Rights, Automation, and the Dehumanization of Lethal Decision-making," *International Review of the Red Cross* 886, no 94 (Summer 2012): 690.
[5] Asaro, *International Review of the Red Cross*, 690.

involvement in lethal decision-making."[6] While these systems may use some level of artificial intelligence (AI)—defined by the Department of Defense (DoD) as "the ability of machines to perform tasks that normally require human intelligence"[7]—that is not the focus of this analysis. The level of artificial intelligence incorporated into the system is irrelevant to the specific ethical questions being asked in this paper.

The DoD Unmanned Systems Integrated Roadmap from 2011 identified four different levels of autonomy which provide a helpful vocabulary framework:

Table 1: Levels of Autonomy

	Name	Description
1	Human Operated	A human operator makes all decisions. The system has no autonomous control of its environment although it may have information-only responses to sensed data.
2	Human Delegated	The vehicle can perform many functions independently of human control when delegated to do so. This level encompasses automatic controls, engine controls, and other low-level automation that must be activated or deactivated by human input and must act in mutual exclusion of human operation.
3	Human Supervised	The system can perform a wide variety of activities when given top-level permissions or direction by a human. Both the human and the system can initiate behaviors based on sensed data, but the system can do so only if within the scope of its currently directed tasks.
4	Fully Autonomous	The system receives goals from humans and translates them into tasks to be performed without human interaction. A human could still enter the loop in an emergency or change the goals, although in practice there may be significant time delays before human intervention occurs.

The first level represents a "human-in-the-loop" system, meaning the human is making all of the substantive decisions and the machine is remote controlled by the human

[6] Asaro, 690. The DoD uses a similar definition in DoD Directive 3000.09: "A weapon system that, once activated, can select and engage targets without further intervention by a human operator. This includes human-supervised autonomous weapon systems that are designed to allow human operators to override operation of the weapon system, but can select and engage targets without further human input after activation." Department of Defense, *Autonomy in Weapons Systems*, DoD Directive 3000.09 (Washington, DC: Department of Defense, 2012), https://www.esd.whs.mil/portals/54/documents/dd/issuances/dodd/300009p.pdf

[7] Department of Defense, *Summary Of The 2018 Department Of Defense Artificial Intelligence Strategy*, (Washington, DC: Department of Defense, 2018), 5, https://media.defense.gov/2019/Feb/12/2002088963/-1/-1/1/SUMMARY-OF-DOD-AI-STRATEGY.PDF

operator.[8] Levels two and three are variations of a "human-on-the-loop" system, where the human operator has pre-programmed the machine to operate self-sufficiently but the human operator stands by on the system with the ability to override any of the decisions made by the autonomous system.[9] Finally, level four represents a "human-out-of-the-loop" system, where the system can operate autonomously and the human operator is not on standby.[10] It should be noted, however, that while the human operator is not on-the-loop, that does not mean that human operators are incapable of regaining control over the system if so required.[11]

Finally, the term non-ethical is used in this paper to mean without an ethical component. This is distinguished from unethical, which means an ethically incorrect decision. Non-ethical means that no ethical decision can or will be made.

III. Ethical Objections to Lethal AWS

Since the birth of the concept of lethal AWS, there has been a backlash against their use. The International Committee for Robot Arms Control was created in 2009 and has been steadily producing academic literature arguing against the employment of lethal AWS ever since.[12] Much of this literature—if even discussed from an ethical perspective—takes as its starting premise that lethal AWS are unethical.[13] When ethical

---

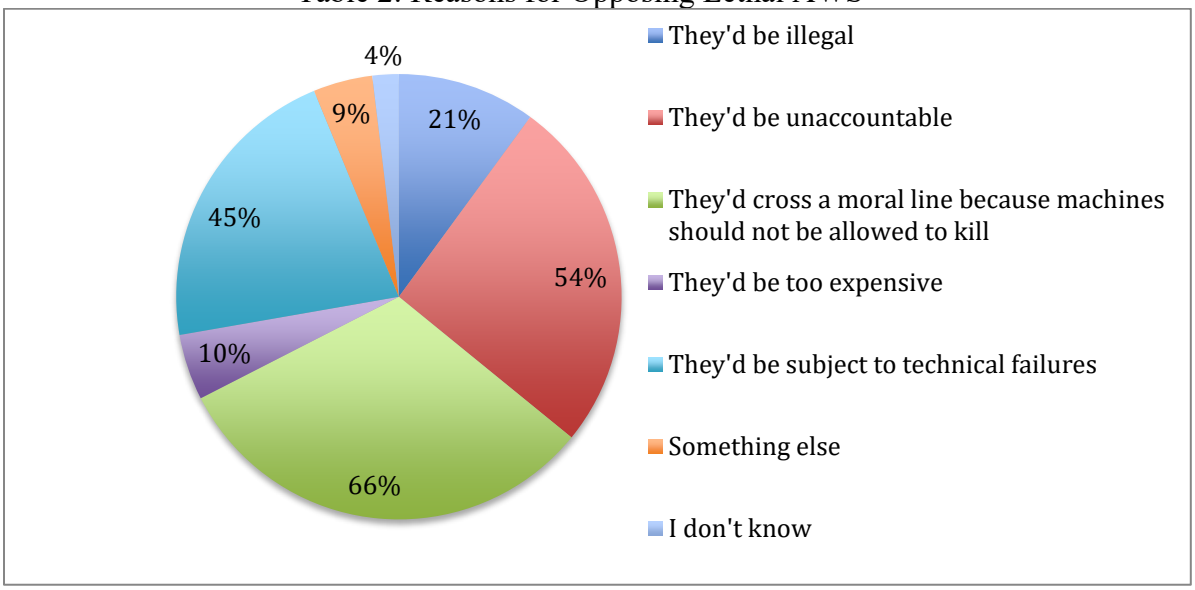
[8] Christian Enemark, *Armed Drones and the Ethics of War* (New York: Routledge, 2014), 100; Alex Leveringhaus, *Ethics and Autonomous Weapons* (London: Palgrave Macmillan, 2016) 3.
[9] Enemark, *Armed Drones and the Ethics of War*, 100; Leveringhaus, *Ethics and Autonomous Weapons*, 3.
[10] Leveringhous, *Ethics and Autonomous Weapons*, 4.
[11] Enemark, *Armed Drones and the Ethics of War*, 100.
[12] "Research," International Committee for Robot Arms Control, accessed May 16, 2019, https://www.icrac.net/research/
[13] It is clear that this is because much of the literature is intended political advocacy and not as ethical analysis; however, that does not change the ethical position being taken.

aspects of the system arise, the discussion often revolves around the ethics of how the AWS is employed, not whether there is an ethically objectionable aspect to the existence of AWS themselves. This section will demonstrate that there is a moral line that is crossed when AWS are given ethical agency for purposes of solving ethical problems.

1. Lethal AWS Cross a Moral Line

The public polling firm Ipsos, sponsored by Human Rights Watch, released a survey in January of 2019 on attitudes towards lethal fully AWS.[14] The survey was focused on two major questions: (1) do you support or oppose lethal AWS; and (2) if not, why not? 61% of the respondents were somewhat or strongly opposed to the use of lethal AWS.[15] The reasons why those 61% were opposed to lethal AWS are outlined in table 2:

Table 2: Reasons for Opposing Lethal AWS



Adapted from Ipsos Press Release. See note 12.

[14] Chris Deeny, *Six in Ten (61%) Respondents Across 26 Countries Oppose the Use of Lethal Autonomous Weapons Systems*, (Washington, D.C.: Ipsos, January 22, 2019), https://www.ipsos.com/sites/default/files/ct/news/documents/2019-01/human-rights-watch-autonomous-weapons-pr-01-22-2019_0.pdf.
[15] Deeny, *Opposition to Lethal AWS.*

The clear majority of participants who opposed lethal AWS opposed them because they believed machines "selecting targets and attacking those targets without human intervention" would cross a moral line.[16] It is this "moral line" that is of particular interest. Statistically, this study shows that a major consideration for individuals who disagree with the concept of lethal AWS is that such a system would cross a moral line; however, this study does not provide any ethical basis for the moral line that is being crossed.[17]

2. Existing Arguments for the Ethical Basis of the Moral Line

Much of literature discussing lethal AWS accepts as true that there is a moral line that should not be crossed; however, this fact is stated without analyzing the underlying ethical basis for the moral line. Robert Sparrow, in his seminal work on ethical responsibility for lethal AWS, devotes only one line to the ethical basis of the moral line: "The thought that a machine might be trusted to make the decision to take a human life is obviously a disturbing one."[18] Sparrow provides no further explanation of why he believes it to be obvious or what ethical construct is being violated; *res ipsa loqutor*—the thing speaks for itself.

Armin Krishnan, in his book *Killer Robots*, devotes an entire chapter to ethical considerations of lethal autonomous robots, with a specific section devoted to automated

---

[16] Deeny, *Opposition to Lethal AWS*.
[17] To be clear, a single study of this type is of limited academic value. Its inclusion here is not to demonstrate the factual efficacy of its results; instead, it is to show, as a general proposition, that some measure of the general public has an ethically based discomfort with the concept of lethal AWS.
[18] Robert Sparrow, "Killer Robots" *Journal of Applied Philosophy 24*, no. 1 (2007): 68.

killing.[19] Krishnan starts the section with the observation: "there are certain things that should only be done by human beings and should never be delegated to machines because machines are in some crucial aspects different from us."[20] This setup, however, is followed by Krishnan outlining various computational differences between humans and computer systems: inability to distinguish simulation from reality; lack of empathy; and no concept of death.[21] While each of these may be true, they do not provide an ethical explanation for the moral line that is crossed with lethal AWS. Krishnan does argue that a robot is not a moral agent, but does so purely within the context of machines being unable to feel remorse or be punished.[22] If the respondents in the study, however, believed that the moral line crossed by lethal AWS was that those systems could not be held responsible, they could have responded with "unaccountability" as their main concern. The moral line, therefore, cannot be merely that mechanical systems are incapable of being morally responsible for their actions.

Peter Asaro, a prominent philosopher of technology and co-founder of the International Committee for Robot Arms Control, identifies two questions in his requirement for human judgment in legal killing. First, *can* a computer make the life and death decisions required for compliance with IHL with a level of performance that is "deemed acceptable?"[23] Second, *ought* a computer make the life and death decisions required for compliance with IHL with a level of performance that is "deemed

---

[19] Armin Krishnan, *Killer Robots: Legality and Ethicality of Autonomous Weapons* (Burlington VT: Ashgate Publishing, 2009) 130-34.
[20] Krishnan, *Killer Robots*, 131.
[21] Krishnan, 131-33.
[22] Krishnan, 132.
[23] Asaro, *International Review of the Red Cross*, 699 (see n. 4).

acceptable?"[24] After posing both questions, Asaro proceeds to focus only on the consequentialist perspective of the former. While noting that computers can comply with the rules of chess, he argues that the rules for IHL require a level of interpretive judgment that is not present in chess.[25] While this may be true, his argument is focused entirely on the consequentialism—is it possible for an AWS to comply with IHL—while ignoring the deontological question posed by his second question.

Asaro never analyzes the ethical basis for his deontological argument. He poses the question ought a "computer, machine or automated process…make these decisions of life or death at all," without every providing any explanation for what it is he finds transgressive about the machine doing so. In other words, Asaro merely poses the question without providing the underlying duty that is violated by allowing lethal AWS to make decisions of life or death.

Alex Leveringhaus, in his book *Ethics and Autonomous Weapons*, comes closest to providing an underlying ethical basis for the moral line.[26] Leveringhaus distinguishes his argument—what he calls the second wave—from the first wave of ethicists who focused on issues of responsibility.[27] Leveringhaus instead focuses his ethical analysis on the idea that there is something "morally valuable" that is lost when human agency is replaced by machine technology.[28] Specifically, Leveringhaus focuses on the inability of machines to know when *not* to pull the trigger even when they are legally authorized to

---

[24] In this analysis, Asaro is mirroring Immanuel Kant's deontological formulation of the categorical imperative. Immanuel Kant, *Foundations of the Metaphysics of Morals*, trans. Lewis White Beck (Upper Saddle River, NJ: Prentice-Hall Inc., 1997) 15-19.
[25] Asaro, *International Review of the Red Cross*, 699.
[26] Leveringhaus, *Ethics and Autonomous Weapons*, 89-117.
[27] Leveringhaus, 87.
[28] Leveringhaus, 91-92.

do so.[29] This argument, however, diagnoses a symptom as the disease. The ability of the lethal AWS to know whether or not to engage at a level consistent with what is allowable under IHL or to restrain itself in accordance with more restrictive ROE can be programmed into the system. The system could then be set for the level of risk acceptable to the decision-maker for the specific conflict: e.g. restrictive ROE for a limited conflict and the entire bounds of IHL for unlimited warfare. While Leveringhaus' ethical argument did not hit the mark, he at least identified that an ethical framework must be constructed to explain the general belief that lethal AWS cross a moral line.

3. Ethical Agency as a Basis of the Moral Line

Armin Krishnan concludes his section on the ethical considerations of automated killing by saying: "in ethical terms there would be no difference between a person killed by a mine or a person killed by an autonomous military robot."[30] This is fundamentally wrong both in terms of substance and ethical understanding. Substantively, a mine is not an autonomous system but an automated system; it has a single response—exploding—that occurs based on a single stimulus—pressure. To be effective, a mine does not require any computing power; it is a pure mechanical response. There is no ethical agency given to the mine to decide when or how to explode. To be clear, there are ethical questions surrounding the use of mines as evident from the creation of the Ottowa Treaty; however, those are ethical questions with how to employ the weapon, not ethical questions with the process by which the weapon itself acts. A lethal AWS, on the other hand, is, by definition, making life and death decisions. Unlike a mine where the ethical

[29] Leveringhaus, 89-94.
[30] Krishnan, *Killer Robots*, 133.

burden is on the human making decisions of employment, with a lethal AWS the machine itself assumes at least some part of the ethical burden.[31]

The Ipsos survey response identified a moral line and amplified that line with the statement "machines should not be allowed to kill." Taken at face value, however, this leads to an absurd result. Most people—setting aside those who have holistic ethical objections to war as a concept—would not have any ethical qualms with a gun being used to kill on a battlefield in time of war; however, a gun is a machine. It seems logical, therefore, that the moral line identified is not related to the use of machines, *per se*, but those machines that are capable of some level of—at least—automation, but most likely autonomy. Similarly, few have ethical qualms with the use of precision-guided munitions within an armed conflict; in fact, arguments have been made for the ethical *requirement* to use precision-guided munitions when available because they lower the risk of civilian casualties.[32] The lack of ethical consternation at the use of precision-guided munitions shows that a system that relies upon automation—and some amounts of autonomy—does not cross a moral line. The moral line, therefore, cannot categorically be that machines are allowed to kill.

The moral line, identified by respondents in the Ipsos survey, is the line of ascribing ethical agency to a non-human. If a review of machine lethality demonstrates that there are no ethical objections to machine lethality with various levels of agency up

---

[31] It is also possible that the individual who programmed the lethal AWS and the individual who chose to employ the lethal AWS would bear some portion of the ethical burden; that said, the decision of when to employ lethal force would be made wholly by the lethal AWS requiring it to bear an ethical burden.

[32] Human Rights Watch, *Off Target: The Conduct of the War and Civilian Casualties in Iraq* (New York: Human Rights Watch, 2003), 15, https://www.hrw.org/reports/2003/usa1203/usa1203.pdf

to lethal AWS, but lethal AWS cross a moral line, the next question must be: what makes lethal AWS ethically different? As shown above, much of the academic literature focuses on the difference in ability to ascribe responsibility between humans and lethal AWS. This, however, is a red herring, as responsibility cannot be ascribed to existing defensive automated weapons systems such as the CWIS; yet, there are no ethical objections to their employment. The difference is not one of responsibility, but one of agency. Decisions of life or death go straight to the heart of ethical thought; namely, how should humans treat each other. Programming a lethal AWS to make an autonomous decision about whom it should kill is ascribing ethical agency to the system. In other words, the problem is not *allowing* machines to kill, but allowing machines to *decide* to kill.

On the other hand, it could be argued that ethical decision-making should be programmed into the lethal AWS, transferring the ethical burden to the software engineer. Ronald Arkin, a leading robot ethicist, makes an argument for programing robots with an "ethical governor."[33] The purpose of this governor is to "conduct an evaluation of the ethical appropriateness of any lethal response that has been produced by the robot architecture prior to its being enacted."[34] Without getting into the technical side of whether or not the software exists to make this possible, this argument has a major underlying flaw; specifically, what is the ethical concept against which the robot's actions are going to be judged?

To explain this question, it is necessary to provide an example. The "trolley problem," is a well-known ethical dilemma in which a runaway trolley is going down the

---

[33] Arkin, *Governing Lethal Behavior*, 127 (see note 3).
[34] Arkin, 127.

tracks out of control. A person at a switch can either do nothing and allow the trolley to hit and kill five people or the person can pull the switch and move the trolley onto a side rail where it will kill one person. If applying consequentialist ethics, the answer is clearly pull the switch; killing one person is a better outcome than killing five. If applying a deontological perspective, the answer is clearly don't pull the switch; any action taken by the person that results in killing another human is a violation of the duty against killing, despite the demonstrably worse outcome.[35] This vignette is given to demonstrate that any ethical governor would have to be programmed not just with a general ethical decision-making capability but with a specific ethical framework to be applied in the situation. In other words, the software engineer would not be building skill into the machine, but building a specific ethical framework into the machine.

Part of what makes ethics such a difficult moral quandary is deciding which ethical construct to apply in any given situation. Ethicists spill much ink to argue for which is the most appropriate ethical framework to apply based on the particular facts of the question, but there is no universally accepted answer. This begs the question, which ethical framework should be programmed into lethal AWS? Is that a decision the software engineer can make? One possible solution would be to program all of the ethical frameworks (e.g. consequentialist, deontological, virtue ethics, etc.) into the lethal AWS.[36] If, however, the system were programmed with all of the various ethical

---

[35] For a fascinating examination of human expectations of robot ethics pertaining to the trolley problem see: Bertram F. Malle et. al., "Sacrifice One for the Good of Many? People Apply Different Moral Norms to Human and Robot Agents," in *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction* (New York: ACM, 2015).

[36] For further reading on consequentialism, see: Jeremy Bentham, *An Introduction to the Principles of Morals and Legislation* (Scotts Valley, CA: CreateSpace, 2017). For

perspectives, this merely pushes the can down the road. There would still have to be ethical agency applied to decide which ethical system to apply whenever an ethical decision arises. In other words, the ethical agency that is inherent to humanity is not present in machines. Whether it cannot or ought not be replicated goes beyond the scope of this paper; suffice it to say that allowing ethical agency in autonomous robots crosses a moral line because ethical reasoning is an inherently human function.

Accepting that there is a moral line crossed when allowing lethal AWS to exercise moral agency puts the current legal literature advocating against lethal AWS in focus. Rather than being an academic body of work based on legal compliance with IHL, the literature is actually premised on an underlying ethical objection to lethal AWS having ethical agency.

IV. Legal Objections to Lethal AWS.

While much of the academic literature on lethal AWS revolves around whether these systems are capable of complying with IHL, a closer look at those legal objections reveals that the objections are not legal, but ethical in nature. If the problem were one purely of specific compliance—that is, ability of a lethal AWS to comply with each rule in a stated rule set—the legal analysis would be relevant. As will be shown, however, the legal analysis focuses not on specific compliance but on the inability of lethal AWS to be able to apply decision-making to legal concepts autonomously. Despite being couched in legal terms, this is not a legal, but an ethical objection. This assessment will focus on the application of lethal AWS to IHL, specifically, the concepts of necessity, distinction and

---

further reading on deontology, see: Immanuel Kant, *Metaphysics of Morals* (Upper Saddle River, NJ: Prentice-Hall, Inc., 1997). For further reading on virtue ethics, see Aristotle, *Nicomachean Ethics*, trans. Robert C. Bartlett and Susan D. Collins (Chicago: University of Chicago Press, 2011).

proportionality. This is not to say that these are the only legal constraints on lethal AWS, but that these are the minimum constraints the system would have to meet to comply with international law.

1. Military Necessity

Military necessity is "the principle that justifies the use of all measures needed to defeat the enemy as quickly and efficiently as possible that are not prohibited by the law of war."[37] Under the United States' understanding of this principle, resort to all means of violence necessary to defeat the enemy as quickly as possible is permissible as long as it is not violating any other rules of IHL.[38] This principle is juxtaposed with the concept of humanity, which "forbids the infliction of suffering, injury, or destruction unnecessary to accomplish a legitimate military purpose."[39]

The argument for the inability of lethal AWS to comply with military necessity hinges on the interaction between necessity and humanity.[40] According to Human Rights Watch, the balancing of military necessity against the requirements for humanity creates a subjective analysis.[41] Essentially, Human Rights Watch argues that the balancing required between military necessity—the measures required to defeat the enemy—on one hand, and humanity—infliction of unnecessary suffering, injury or damage—on the

---

[37] Department of Defense, *Department of Defense Law of War Manual* (Washington DC: Department of Defense, 2016), sec. 2.2, https://dod.defense.gov/Portals/1/Documents/pubs/DoD%20Law%20of%20War%20Manual%20-%20June%202015%20Updated%20Dec%202016.pdf?ver=2016-12-13-172036-190.
[38] Department of Defense, *Law of War Manual*, sec. 2.2
[39] Department of Defense, sec. 2.3
[40] Human Rights Watch, *Losing Humanity*, 34.
[41] Human Rights Watch, 34.

other, relies on the "context-dependent, value-based judgment of a commander."[42] While not explicitly stated as part of the legal argument, the reader is meant to infer that this type of decision-making is not appropriate for a lethal AWS. In other words, the problem is not with the lethal AWS's ability to comply with the rules, it is with that system's inability to apply an ethical framework to balance between two applicable principles of IHL.

2. Distinction

Distinction, a fundamental principle of international law that has been identified as "intransgressable" by the International Court of Justice,[43] holds that civilians—unlawful targets—must be distinguished from combatants—lawful targets.[44] When dealing with state-on-state conflict, this process is relatively straightforward. Those individuals who are identified as being a member of the armed forces of the opponent's armed forces are valid legal targets.[45] As members of armed forces are required to wear identifying markings, these individuals are usually easily distinguishable by their uniforms.

The legal analysis, however, becomes more complicated when dealing with non-state threats. For example, civilians that are directly participating in hostilities are valid

---

[42] Human Rights Watch, *Losing Humanity*, 33 (quoting Benjamin Kastan, "AWS: A Coming Legal "Singularity"?" *University of Illinois Journal of Law, Technology and Policy* 2013, no 45 (Spring 2013) 59.
[43] Legality of the Threat or Use of Nuclear Weapons, Advisory Opinion, 1996 I.C.J. 226 (July 8), 78.
[44] Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts (Protocol I) art 38, Jun. 8 1977, 1125 U.N.T.S. 3 [hereinafter Additional Protocol I].
[45] Allowing for exceptions such as chaplains and medical professionals in accordance with Geneva Convention for the Amelioration of the Condition of the Wounded and Sick in Armed Forces in the Field art. 24, Aug. 12 1949, 6 U.S.T. 3114, 75 U.N.T.S. 31 [hereinafter Geneva Convention I].

targets under the rule of distinction. Three guidelines have been set forth for identifying if a civilian is directly participating in hostilities: (1) threshold of harm; (2) direct causation; and (3) belligerent nexus.[46] As one author put it, these guidelines are "to help a moral agent through multiple layers of interpretation and judgment."[47] Put another way, being able to distinguish between a civilian who is protected and civilian who is directly participating in hostilities is based on "situational awareness and having a theory of mind, that is, understanding someone else's intentions and predicting their likely behavior in a particular situation."[48] This "theory of mind" is, specifically, the ability to apply ethically based decision-making.

While couched in a legal critique of whether lethal AWS can comply with IHL, the underlying objection being made by these authors is not one of legal compliance, but one of ethical competence. The argument presented is not that lethal autonomous weapon systems do not comply with the principle of distinction under IHL; instead, it is that lethal AWS are *incapable* of complying with distinction because it requires ethical decision-making.

3. Proportionality

Proportionality prohibits "an attack which may be expected to cause incidental loss of civilian life, injury to civilians, damage to civilian objects, or a combination thereof, which would be excessive in relation to the concrete and direct military

---

[46] Nils Mezler, *Interpretive Guidance on the Notion of Direct Participation in Hostilities Under International Humanitarian Law* (Geneva: International Committee for the Red Cross, 2009), 47-64, https://www.icrc.org/en/doc/assets/files/other/icrc-002-0990.pdf
[47] Asaro, *International Review of the Red Cross*, 698.
[48] Noel Sharkey, "Killing Made Easy: From Joysticks to Politics," in *Robot Ethics: The Ethical and Social Implications of Robotics*, ed. Patrick Lin, Keith Abney, and George A. Bekey (Cambridge, MA: The MIT Press, 2012), 118.

advantage anticipated."[49] To be clear, proportionality does not require a blanket prohibition on damage to civilian objects or loss of civilian life; instead, that must be balanced against the weight of the military necessity.[50] As the import of the military target rises, the amount of harm that can be caused to civilians and civilian objects rises commensurately.

The legal argument against lethal AWS being able to comply with proportionality is focused on complexity of application.[51] Determination of proportionality is abstract, "not easily quantified, and highly relative to specific contexts and subjective estimates of value."[52] This abstract, fact-specific determination is made by commanders in the battlespace that the attack will not result in excessive civilian casualties as compared to the weight of the necessity of attacking that target. Human Rights Watch specifies that a robot could not be programmed to "duplicate the psychological processes in human judgment that are necessary to assess proportionality."[53]

While couched in legalistic terms, the fundamental objection raised is not compliance based but that the calculations required to balance the military necessity of the object with the cost of civilian lives killed requires ethical agency. In order for a lethal AWS to be able to weigh the military necessity of a particular target against the cost of harm to humans, that system must be able to conduct ethical decision-making.

---

[49] Additional Protocol I art 51(5)(b).
[50] Department of Defense, *Law of War Manual*, sec. 2.4.1.2.
[51] Human Rights Watch, *Losing Humanity*, 32.
[52] Peter Asaro, "Modeling the Moral User," *IEEE Technology and Society Magazine*, 28, no.1 (Spring 2009) 21. http://peterasaro.org/writing/Asaro%20Modeling%20Moral%20User.pdf.
[53] Human Rights Watch, *Losing Humanity*, 33.

Although not specifically stated in the legal critique, it is the fact that machines would be required to do ethical decision-making that underlies the entire legal critique.

4. Conclusion

Despite the use of academic legal language to offer a critique of lethal AWS, a review of the arguments shows that they are all based on the premise that lethal AWS should not be allowed ethical agency. These critics then take the next step to argue that if some decision-making to comply with IHL requires ethical agency and autonomous AWS should not have ethical agency, then lethal AWS cannot comply with IHL. This argument is based on the logical flaw that if one aspect of IHL cannot be complied with, all of IHL cannot be complied with. Instead, it is necessary to identify how much—if any—of the targeting process under IHL a lethal AWS could perform without requiring any ethical decision making on the part of the autonomous system. Put another way, accepting as a constraint that lethal AWS are incapable of ethical agency, can they still provide military benefit on the battlefield?

V. Applying Non-Ethical Lethal AWS to Targeting

While many have argued that because some aspects of IHL require ethical agency, and lethal AWS cannot make decisions requiring ethical agency, lethal AWS are, *per se*, in violation of IHL. Instead, it is necessary to identify how much—if any—of the targeting process a non-ethical lethal AWS could perform without requiring it to make any ethical decisions. Instead of throwing out the baby with the bathwater, it is necessary to walk through the legal targeting process to identify where there are decisions that require ethical agency and, therefore, cannot be done by lethal AWS.

1. *Targeting Process*

The targeting process is the process by which IHL is applied to a list of targets to ensure those targets comply with international law. For purposes of this section, a target is "an entity or object that performs a function for the threat considered for possible engagement or other action."[54] The principles of IHL can be applied sequentially, to ensure compliance with each step. While there is more than one way to sequence and group these principles, for the purposes of this article, the sequence will be: (1) Military Necessity; (2) Distinction; (3) Proportionality; and (4) Humanity.[55]

2. Military Necessity

The initial question for military necessity is: does the immediate action assist towards the object of the war? Many legal scholars have pointed out that military necessity is both difficult to define and apply.[56] On its face, this is true; military necessity can conceptually be balanced against humanity as two poles in an ethical decision. On the one hand, military necessity allows all means of violence necessary to defeat the enemy as quickly as possible, without taking any action that is otherwise prohibited by IHL; on the other, humanity forbids the unnecessary infliction of suffering, injury, or destruction.[57]

This, however, is an unnecessarily complicated way of construing the principles; instead, it is helpful to separate military necessity from humanity. This is not to say that

---

[54] Joint Chiefs of Staff, *DoD Dictionary of Military and Associated Terms* (Washington, DC: Joint Chiefs of Staff, 2019), https://www.jcs.mil/Portals/36/Documents/Doctrine/pubs/dictionary.pdf.
[55] For an alternate grouping option see LCDR Luke Whittimore, "Proportionality Decision Making in Targeting: Heuristics, Cognitive Biases, and the Law," *Harvard National Security Journal* 7, no. 577 (2016): 588-89.
[56] Department of Defense, *Law of War Manual*, sec. 2.2.3
[57] Department of Defense, secs. 2.2 and 2.3

humanity should be excluded from the targeting process; it should be included as a separate logical step—as it is at the end of this targeting process—from military necessity. By simplifying and breaking the questions down into their constituent parts, military necessity can be boiled down to: "does the action have the objective of weakening the enemy forces?[58]

In addition to this simplified understanding of military necessity as divorced from humanity, there are actions that have been recognized as generally inherently militarily necessary.[59] These actions include things such as targeting enemy combatants or interning enemy POWs.[60] These actions have—elsewhere in the *lex scripta* of IHL—defined rules of constraint or restraint. For those areas of IHL where specific rules apply, there is no need to resort to general principles.[61] In other words, wherever there is a specific rule in IHL allowing for a particular action, that rule renders the requirement to assess military necessity moot.

As applied to lethal AWS, the question "does the action have the objective of weakening the enemy forces" does not require ethical agency to solve. This question is fact-based and not open to interpretation. For example, an attack on a cruise ship carrying civilian passengers has no military necessity. An attack on a hydroelectric dam, may have military necessity—e.g. if it were generating electricity being utilized by the military forces. This distinction between targets that are militarily necessary and those that are not is a fact-based assessment; it does not require any ethical agency. That said, military necessity is only the first step of targeting and merely assessing that a target

---

[58] Whittimore, *Harvard National Security Journal*, 588.
[59] Department of Defense, *Law of War Manual*, sec. 2.2.3.2.
[60] Department of Defense, sec. 2.2.3.2.
[61] Department of Defense, sec. 2.2.3.2

meets this requirement is not *carte blanche* to target it; instead, IHL requires moving to the next step.

3. Distinction

i. Distinction of Combatants and Civilians

The principle of distinction—sometimes referred to as discrimination—requires any party to a conflict to differentiate between the armed forces and civilian population of their opponent, with only the former being valid legal targets.[62] Conceptually, distinction is broken into two categories: (1) discrimination pertaining to the enemy; and (2) discrimination pertaining to one's own forces.[63] While the majority of this analysis will focus on the former, as it is the category relevant to targeting, the requirement to differentiate between one's own combatants and civilians is also relevant to the functionality of lethal AWS. When dealing with human targets, the targets will fall into one of three categories: (1) members of the armed forces; (2) civilians; or (3) members of an organized armed group.

Members of armed forces present the most straightforward analysis for non-ethical lethal AWS. Members of an armed force are required to differentiate themselves from civilians, in accordance with the concept of discrimination of one's own forces.[64] This means that members of an armed force should have distinctive uniforms, or other such visible distinguishing marks. Similarly, categories of persons who are in the armed forces, but are protected—such as medical and religious personnel—are also required to

[62] Department of Defense, sec. 2.5.
[63] Department of Defense, sec. 2.5.
[64] Department of Defense, sec. 2.5.3.1.

wear distinguishing emblems.[65] As a result, there should be no ambiguity for non-ethical lethal AWS to correctly identify both those members of the opposing armed force and those members of the opposing armed force who are entitled to protected status under IHL.

The second category of potential targets is civilians directly participating in hostilities. Most of the legal arguments that lethal AWS are incapable of complying with the principle of distinction focus on this particular category of distinction.[66] Under this concept, civilians—by actively participating in the conflict—forfeit the protections afforded to civilians under IHL.[67] Immediately, it can be seen how difficult this distinction can be in practice. For example, if a civilian farmer picks up a discarded weapon, takes two shots at an opposing force, drops the weapon and returns to farming, can he be targeted? According to the DoD Law of War Manual, the answer is highly contextual and will incorporate the following non-dispositive considerations: amount of harm caused by the act; connection between the act and hostilities; purpose for the underlying act; military significance of the underlying act; and whether the act was traditionally military in nature.[68] This type of weighing analysis—particularly as pertains to discerning intent of the actor—requires the exact type of ethical agency that should not be given to lethal AWS; as such, this category of target would not be targetable by those systems.

A third category of potential targets is organized armed groups. The Geneva Convention identifies that not all combatants are members of the armed forces, but for

---

[65] Geneva Convention I, art 40.
[66] Human Rights Watch, *Losing Humanity*, 30.
[67] Department of Defense, *Law of War Manual*, sec. 5.8.
[68] Department of Defense, sec. 5.8.3.

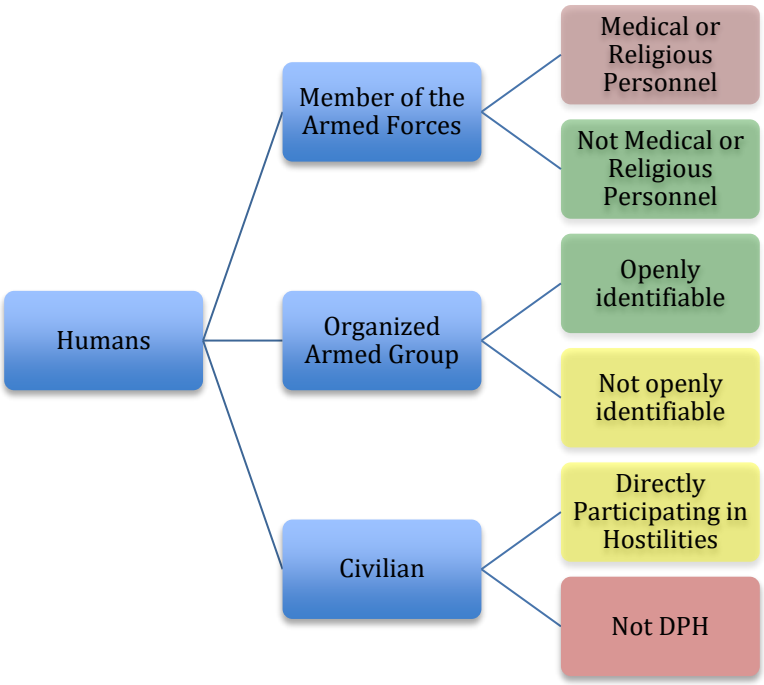
purposes of being given all the rights of combatants, there are specific criteria that an organized armed group must meet. In particular, an organized armed group must: have a commander with responsibility for subordinates; have a fixed distinctive symbol; carry arms openly and conduct their operations in accordance with the laws and customs of war.[69] While these criteria were designed to differentiate who is entitled to combatant status, they were written from the point of view of who is entitled to the benefits associated with combatant status. It is only the second and third requirements that are necessary for treating an organized armed group as combatants for targeting purposes. Otherwise, an organized armed group would be able to argue that they are not targetable because they failed to support the laws and customs of war, a perverse incentive. The fact, however, that the members of organized armed groups are required to both carry their arms openly and have distinctive emblems makes their targeting merely an issue of positive identification not requiring ethical agency. If, however, the organized armed group is not wearing distinctive symbols or carrying their arms openly, for purposes of the lethal AWS the members of that group would not be targetable because they would be indistinguishable from civilians.

In sum, targeting of humans by non-ethical lethal AWS is not, *per se*, prohibited under IHL. Without being required to employ ethical agency, lethal AWS can target both members of the opposing armed force and members of organized armed groups that are complying with the requirements for combatant status. Members of organized armed groups who are not complying with the requirements for combatant status and civilians directly participating in hostilities, while they may be targetable under IHL, would not be

[69] Geneva Convention I art. 4(2)

targetable with a non-ethical lethal AWS. Civilians, medical and religious personnel are not targetable under IHL. A graphical representation of this flow chart is represented in Table 3, with illegal targets under IHL in red, legal targets under IHL that are targetable by non-ethical AWS in green, and legal targets under IHL but not targetable by non-ethical AWS in yellow.

Table 3: Human Targeting Distinction Flow Chart



The red boxes are targets that would be illegal under IHL. The green and yellow boxes are permissible targets under IHL; however, the yellow boxes are ethically impermissible targets for lethal AWS.

ii. Distinction of Military and Civilian Objects

The principle of distinction applies not only to targeting of human beings, but to targeting of objects as well. When targeting objects, the requirement under IHL is that the objects be a military objective that makes an effective contribution to the enemy's military action, the destruction of which would offer the attacker a definite military

advantage.[70] In practice, confirmation under one of the rules usually means confirmation under the other. In other words, having an effective contribution to an enemy's military action is often proof that destruction of the object would give the attacker a definite military advantage and vice versa. Objects qualify as making an effective contribution under one of four possible bases: (1) nature; (2) location; (3) purpose; or (4) use.[71]

Nature pertains to the intrinsic character of the object, such as military equipment.[72] Any type of military equipment—whether or not it is specifically for combat purposes—aids the enemy in conducting hostilities; therefore, military equipment is always targetable. For purposes of non-ethical lethal AWS, this category of object would be targetable because its benefit to the enemy would be obvious on its face. A list of items with intrinsic military value could be programmed into the machine and no ethical agency would be required.

Location pertains to targeting a particular geographic locale rather than an object at that locale.[73] For example, if a particular area is known as a smuggling route for contraband weapons to the enemy, that location would be targetable to prevent it from continuing to be used.[74] For purposes of non-ethical lethal AWS, this type of analysis would require ethical agency. Understanding where to target a location requires complex synthesis of intelligence information with pattern of life information to form an educated guess as to which location would give military benefit. This information could be

---

[70] Michael Schmitt, "Targeting in Operational Law," in *The Handbook of the International Law of Military Operations*, ed. Terry Gill and Dieter Fleck, 2nd ed. (Oxford: Oxford University Press, 2015) 278.
[71] Schmitt, "Targeting in Operational Law," 279-80.
[72] Schmitt, 279.
[73] Schmitt, 279.
[74] Schmitt, 279.

calculated as part of the intelligence function of a military command and fed to the lethal AWS by a man-on-the-loop or a man-in-the-loop, but it could not make those calculations without ethical agency of its own.

Purpose pertains to the future use of an object.[75] A particular civilian object—if being prepared for military use—need not be in military use yet in order to be targeted.[76] In other words, a civilian object is targetable when it becomes clear there is a plan for military use of that object; actual military use is not required. As with location, above, if there is specific intelligence that an object will be used for a military purpose that is passed to a non-ethical lethal AWS through a man-on-the-loop or a man-in-the-loop, it would be able to target that object. Without specific intelligence as to future use, discerning possible future use for a dual use object requires ethical agency to weigh the likelihood possible future use against the possible harm caused by the attack.

Use pertains to an object's current employment.[77] Objects that have a civilian function, when used by the military, become targetable objects. For example, a civilian bridge is not targetable, unless it is being used by the military to move troops and supplies. There is a strict presumption that a civilian object retains its civilian character and: "An object can only be attacked if, based on all readily available information, a reasonable attacker would conclude that it is being used militarily.[78] On its face, this would appear to be a determination that requires ethical agency; after all, what is "reasonableness" if not the application of ethical decision-making? That said, this reasonableness test is only necessary when the status of the use is unclear. If, for

---

[75] Schmitt, 279.
[76] Schmitt, 279.
[77] Schmitt, 280.
[78] Schmitt, 280.

example, there is a civilian airplane transporting the military, that is purely a fact-based analysis; there is no requirement to apply ethical decision-making. For purposes of non-ethical lethal AWS, the analysis would be limited to actual use at the time the weapons system was sensing the object. This would obviate the need to apply any reasonableness test as to whether the object could be targeted.

In sum, for targeting objects, a non-ethical lethal AWS would not require the ability to employ ethical decision-making for targeting objects that by their nature or use were military objects. Objects, however, that were military objects by their location or purpose would not be targetable by a non-ethical lethal AWS, unless that determination was made by humans and given to the system.

4. Proportionality

Proportionality is a recognition that in time of war, harm to civilians and civilian property is unavoidable, but that harm must be proportional to the military benefit gained from the attack.[79] To be clear, this is not a prohibition on incidental damage from attacks; rather, it is an affirmation that that incidental damage should not be excessive in relation to the military benefit.[80] This analysis is applied at the time the attack is planned. In other words, the expected harm, as was known at the time of planning, is compared to the expected military benefit at the time of planning. If the planning information turns out to be incorrect—for example, the target was not actually in the believed location—that does not invalidate the proportionality assessment as made during planning.[81]

---

[79] Department of Defense, *Law of War Manual*, sec. 2.4.1.2.
[80] Department of Defense, sec. 2.4.1.2.
[81] Schmitt, "Targeting in Operational Law," 284.

There are two parts to the balancing test, each with a different ethical agency required to calculate: (1) the harm caused; and (2) the military advantage gained. On its face, the harm caused analysis appears to require ethical agency. It requires an ethical assessment of the worth of the person or object being damaged. As practiced, however, the United States military uses a process called "collateral damage assessment methodology" (CDEM) to assess the likelihood of damage to civilians or civilian objects near the target.[82] The CDEM is a calculation that takes into account: "the precision of a weapon, its blast effect, attack tactics, the probability of civilian presence in structures near the target, and the composition of structures to estimate the number of civilian casualties likely to be caused during an attack."[83] In this sense, the calculation of harm caused is a purely rational calculation; information is plugged into an equation and the level of harm is spit out. As such, a non-ethical lethal AWS could be able to complete this aspect of proportionality without any ethical agency.[84]

The military advantage gained from any particular attack is a contextual analysis.[85] While the level of force may be easily distinguishable, changes in time or space can greatly affect military advantage. For example, attacking a mine laying ship has a higher value prior to the ship laying mines than after the ship has already laid mines. That same mine laying ship, if it is 300nm from the area of conflict, has less military advantage than one that is 10nm away. As such, determination of military

---

[82] Michael Schmitt, "Autonomous Weapon Systems and International Humanitarian Law: A Reply to the Critics" *Harvard National Security Journal* 4, no. 1 (2013): 19.
[83] Schmitt, "Autonomous Weapon Systems," 19.
[84] Schmitt, 20 ("There is no question that autonomous weapon systems could be programmed to perform CDEM-like analyses to determine the likelihood of harm to civilians in the target area.")
[85] Schmitt, 20.

advantage is a highly subjective analysis that requires ethical agency in order to determine.

While a non-ethical lethal AWS is only capable of completing half of the required analysis for determining proportionality, this does not mean it is incapable of complying with IHL for targeting. Proportionality requires these two assessments to be determined in relation to each other. Furthermore, a non-ethical lethal AWS could determine that military advantage can be gained from targeting a specific object, just not how to weight that particular advantage. By setting the CDEM threshold very low—ensuring that there would be very little harm incurred to civilians or civilian objects—non-ethical lethal AWS would only have to identify if there was some military advantage gained. Striking a target with even a low amount of military advantage would comply with the principle of proportionality as long as the CDEM stayed very low. For example, if an enemy had a military communications relay station located in a remote desert, the chance of any collateral damage is very low if not non-existent; therefore, the fact that the relay station is a military target would make it targetable without requiring ethical decision-making. This limits the use of non-lethal AWS to areas where they would incur very low casualty rates incident to lawful targeting.

5. Humanity

The principle of humanity prohibits inflicting suffering, injury or destruction without a legitimate military purpose.[86] Humanity requires that weapons used do not cause superfluous injury or unnecessary suffering.[87] Humanity is also the basis for several IHL *lex scriptura* rules, such as: protections for civilian property; protection for

[86] Department of Defense, *Law of War Manual*, sec. 2.3.
[87] Whittimore, "Proportionality Decision Making in Targeting," 598.

medical personnel; protection for religious personnel; and prohibition on weapons designed to cause superfluous injury.[88] In particular, humanity also provides protection for combatants who are *hors de combat*—i.e. injured to the point of being unable to continue fighting.

Under Article 36 of Additional Protocol I, all new weapons must be reviewed to ensure that they do not cause superfluous injury or unnecessary suffering.[89] As lethal AWS do not bring any new weapons to the battlefield—they will employ the same weapons as used by humans—the only question is if having a robot pull the trigger causes superfluous injury or unnecessary suffering. From the point of view of the target, the injury or suffering caused is not affected in any way by whether the trigger was pulled by a human or a robot; therefore, the only relevant analysis is on the weapon itself. This analysis would happen well before the weapon was ever employed on an autonomous system.

Combatants who are *hors de combat*, while a difficult problem for human combatants would be easier for a lethal AWS. Unlike humans, autonomous systems would have sensors that provide significantly more information about the status of the combatant to identify if they are, in fact, injured. Often, injuries are not immediately visible to the human eye, but advanced sensor technology may give the lethal AWS the ability to diagnose injury better than a human counterpart. Furthermore, unlike when a human is approaching an injured combatant, the lethal AWS does not have to worry about its own mortality; therefore, the system could be programmed to prioritize its own destruction over the possibility of attacking an injured enemy combatant. This ability to

---

[88] Department of Defense, *Law of War Manual*, sec. 2.3.2.
[89] Additional Protocol I art. 36

assume a higher level of risk protects the principle of humanity better than a human would be able to do in the same situation.

6. Utility of Non-Ethical Autonomous Weapons Systems

In sum, a non-ethical lethal AWS would have the ability to comply with IHL in a limited set of circumstances. Specifically, non-ethical lethal AWS would only be able to target members of the opposing military or organized armed groups who could be openly identified. Furthermore, non-ethical lethal AWS would have to have a very low threshold for incidental harm to civilians or civilian objects to allow for its inability to assess weight to the value of the military target. Non-ethical lethal AWS would not have been much use over the past eighteen years in a threat environment characterized by civilians directly participating in hostilities and high levels of possible civilian casualties. That said, the threat environment is changing. With the rise of near-peer competitors in China and Russia, the United States is beginning to change its focus away from violent extremist organizations to great power competition. In great power competition, there is more possible use for non-ethical lethal AWS, particularly, in the maritime environment. Unlike the land, where the likelihood of civilian casualties is very high, in the maritime environment, attacking an opposing naval vessel—assuming the correct vessel is hit—will presumably result in no civilian casualties. Furthermore, the profile of foreign military vessels is easily distinguishable, making positive identification reliable for non-ethical lethal AWS. Given the limitations on targeting imposed by preventing lethal AWS from having ethical agency, maritime great power conflict is the type of conflict that would most take advantage of the targeting capabilities of the platform.

VI. Conclusion

While they have not yet been fielded by the United States military, the technology required to field lethal AWS is not far off. Prior to their arrival, it is in all of our best interest to review how they will be employed and ensure that they are not just technologically capable but ethically acceptable. A study of the general population found that a vast majority who disapproved of lethal autonomous robots did so because they believed lethal AWS crossed a moral line. While the existing ethical literature did not provide a clear answer to where that moral line resides, this paper presented the view that the moral line consists of giving autonomous robots ethical agency. This was reinforced by a review of the anti-lethal AWS legal literature that was actually based on the ethical objection that machines should not be allowed ethical agency.

Accepting as a limiting factor that lethal AWS should not be allowed to have ethical agency, this paper sought to define if they can still function within IHL during the targeting process. In sum, yes, with caveats. Non-ethical lethal autonomous robots can target members of armed forces and members of organized armed groups and comply with the requirements of proportionality, but only in environments where it is possible to ensure very low levels of harm caused to civilians or civilian objects. Given these restrictions, non-ethical lethal autonomous robots will be most useful in a peer-on-peer conflict of the future, particularly a maritime conflict. As such, it is in the interest of the Department of Defense, including the United States Navy, to continue research and development of non-ethical lethal AWS.

BIBLIOGRAPHY

Aristotle. *Nicomachean Ethics*. Translated by Robert C. Bartlett and Susan D. Collins. Chicago: University of Chicago Press, 2011.

Arkin, Ronald. *Governing Lethal Behavior in Autonomous Robots.* Boca Raton, FL: Chapman and Hall, 2009.

Asaro, Peter. "Modeling the Moral User." *IEEE Technology and Society Magazine* 28, no. 1 (Spring 2009): 20-24. http://peterasaro.org/writing/Asaro%20Modeling%20Moral%20User.pdf.

———. "On Banning Autonomous Weapon Systems: Human Rights, Automation, and the Dehumanization of Lethal Decision-making." *International Review of the Red Cross* 886, no. 94 (Summer 2012): 687-709.

Bentham, Jeremy. *An Introduction to the Principles of Morals and Legislation.* Scotts Valley, CA: CreateSpace, 2017.

Deeny, Chris. *Six in Ten (61%) Respondents Across 26 Countries Oppose the Use of Lethal Autonomous Weapons Systems*. Washington, D.C.: Ipsos, January 22, 2019. https://www.ipsos.com/sites/default/files/ct/news/documents/2019-01/human-rights-watch-autonomous-weapons-pr-01-22-2019_0.pdf.

Department of Defense. *Autonomy in Weapons Systems*. DoD Directive 3000.09. Washington, DC: Department of Defense, 2012. https://www.esd.whs.mil/portals/54/documents/dd/issuances/dodd/300009p.pdf

———. *Department of Defense Law of War Manual.* Washington DC: Department of Defense, 2016. sec. 2.2. https://dod.defense.gov/Portals/1/Documents/pubs/DoD%20Law%20of%20War%20Manual%20-%20June%202015%20Updated%20Dec%202016.pdf?ver=2016-12-13-172036-190.

———. *Summary of The 2018 Department of Defense Artificial Intelligence Strategy*. Washington, DC: Department of Defense, 2018. https://media.defense.gov/2019/Feb/12/2002088963/-1/-1/1/SUMMARY-OF-DOD-AI-STRATEGY.PDF

Enemark, Christian. *Armed Drones and the Ethics of War.* New York: Routledge, 2014.

Fleichman, William M. "Just Say "No!" to Lethal Autonomous Robotic Weapons." *Journal of Information Communication and Ethics in Society* 13, no 3/4 (2015): 299-313.

Geneva Convention for the Amelioration of the Condition of the Wounded and Sick in Armed Forces in the Field art. 24, Aug. 12 1949, 6 U.S.T. 3114, 75 U.N.T.S. 31.

Human Rights Watch, *Off Target: The Conduct of the War and Civilian Casualties in Iraq*. New York: Human Rights Watch, 2003. https://www.hrw.org/reports/2003/usa1203/usa1203.pdf

Human Rights Watch and International Human Rights Clinic at Harvard Law School. *Losing Humanity: The Case Against Killer Robots.* New York: Human Rights Watch, November 2012. https://www.hrw.org/sites/default/files/reports/arms1112ForUpload_0_0.pdf.

International Committee for Robot Arms Control. "Research." Accessed May 16, 2019. https://www.icrac.net/research/

Joint Chiefs of Staff. *DoD Dictionary of Military and Associated Terms.* Washington, DC: Joint Chiefs of Staff, 2019. https://www.jcs.mil/Portals/36/Documents/Doctrine/pubs/dictionary.pdf.

Kant, Immanuel. *Metaphysics of Morals*. Upper Saddle River, NJ: Prentice-Hall, 1997.

Kastan, Benjamin. "AWS: A Coming Legal "Singularity"?" *University of Illinois Journal of Law, Technology and Policy* 2013, no 45 (Spring, 2013): 45-82.

Krishnan, Armin. *Killer Robots: Legality and Ethicality of Autonomous Weapons.* Burlington, VT: Ashgate Publishing, 2009.

Legality of the Threat or Use of Nuclear Weapons, Advisory Opinion, 1996 I.C.J. 226 (July 8).

Leveringhaus, Alex. *Ethics and Autonomous Weapons.* London: Palgrave Macmillan, 2016.

Lucas, George. *Military Ethics: What Everyone Needs to Know*. Oxford: Oxford University Press, 2016.

Malle., Bertram F., Matthias Scheutz, Thomas Arnold, John Voiklis and Corey Cusimano. "Sacrifice One for the Good of Many? People Apply Different Moral Norms to Human and Robot Agents." in *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*. New York: ACM, 2015.

Mezler, Nils. *Interpretive Guidance on the Notion of Direct Participation in Hostilities Under International Humanitarian Law.* Geneva: International Committee for the Red Cross, 2009. https://www.icrc.org/en/doc/assets/files/other/icrc-002-0990.pdf

Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts (Protocol I) art 38, Jun. 8 1977, 1125 U.N.T.S. 3.

Santoni de Sio, Filippo and Jeroen van den Hoven. "Meaningful Human Control over Autonomous Systems: A Philosophical Account." *Frontiers in Robotics and AI* 5, no. 15 (February, 2018): 1-14.

Schmitt, Michael. "Autonomous Weapon Systems and International Humanitarian Law: A Reply to the Critics." *Harvard National Security Journal* 4, no. 1 (2013): 1-37.

———. "Targeting in Operational Law." in *The Handbook of the International Law of Military Operations*. Edited by Terry Gill and Dieter Fleck. 2nd Edition. Oxford: Oxford University Press, 2015.

Singer, P. W. *Wired for War: The Robotics Revolution and Conflict in the Twenty-first Century*. New York: The Penguin Press, 2009.

Sharkey, Noel. "Automating Warfare: Lessons Learned from Drones." *Journal of Law Information and Science* 21, no. 140 (2011/2012): 140-54.

Sharkey, Noel. "Killing Made Easy: From Joysticks to Politics." in *Robot Ethics: The Ethical and Social Implications of Robotics*. Edited by Patrick Lin, Keith Abney, and George A. Bekey. Cambridge, MA: The MIT Press, 2012.

Sparrow, Robert. "Killer Robots." *Journal of Applied Philosophy 24*, no. 1 (2007): 62-77.

Verdiesen, Ilse. "Agency Perception and Moral Values Related to Autonomous Weapons: An Empirical Study Using the Value-Sensitive Design Approach." Master's thesis, Delft University of Technology, 2017. http://resolver.tudelft.nl/uuid:7cc28c2e-69d9-45f3-9c87-51e8281c32b0

Wallach, Wendell and Peter Asaro, ed. *Machine Ethics and Robot Ethics*. New York: Routledge, 2017.

Whittimore, LCDR Luke. "Proportionality Decision Making in Targeting: Heuristics, Cognitive Biases, and the Law." *Harvard National Security Journal* 7, no. 577 (2016): 581-636.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
30-05-2019	FINAL	

4. TITLE AND SUBTITLE

Lethal Autonomous Weapons Systems: Can Targeting Occur Without Ethical Decision-Making?

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

LCDR Andrew M. Coffin, JAGC, USN

Paper Advisor (if Any):

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Ethics in Emerging Military Technology
Naval War College
686 Cushing Road
Newport, RI 02841-1207

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Distribution Statement A: Approved for public release; Distribution is unlimited.

13. SUPPLEMENTARY NOTES A paper submitted to the faculty of the NWC in partial satisfaction of the requirements of the Ethics in Emerging Military Technology Graduate Certificate. The contents of this paper reflect my own personal views and are not necessarily endorsed by the NWC or the Department of the Navy.

14. ABSTRACT

The rise in capability of lethal autonomous weapons systems has led to a backlash against their use on both legal and ethical grounds. While the academic literature evinces a general ethical objection to the employment of lethal autonomous weapons, the specific nature of that objection is not clearly explored. Opponents agree that the use of lethal autonomous weapons crosses a moral line; however, they do not elucidate where that line is drawn. This paper argues that the moral line is crossed when lethal autonomous weapons systems are given ethical agency to make life or death decisions free of human input. Furthermore, the legal objections raised to lethal autonomous weapons systems are not inherently legal; instead, they are based on this ethical objection applied to the legal context. Given this underlying objection to lethal autonomous weapons systems being given moral agency, this paper analyzes whether non-ethical lethal autonomous weapons systems—i.e. systems that are not allowed to make any ethical decisions—can comply with international humanitarian law in a way that is still militarily useful. It concludes that non-ethical lethal autonomous weapons systems would be able to comply with international humanitarian law in specific situations, specifically, in near-peer—particularly near-peer maritime—combat.

15. SUBJECT TERMS

Lethal Autonomous Weapons Systems, Ethics, International Law, Law of Armed Conflict

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Chairman, EEMT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED		32	**19b. TELEPHONE NUMBER** *(include area code)* 401-841-3556

Standard Form 298 (Rev. 8-98)