1-1-2014

Autonomous Weapons Systems: Taking the Human out of International Humanitarian Law

James Foy

Follow this and additional works at: <https://digitalcommons.schulichlaw.dal.ca/djls>

This work is licensed under a [Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 License](#).

Recommended Citation

James Foy, "Autonomous Weapons Systems: Taking the Human out of International Humanitarian Law" (2014) 23 Dal J Leg Stud 47.

This Article is brought to you for free and open access by the Journals at Schulich Law Scholars. It has been accepted for inclusion in Dalhousie Journal of Legal Studies by an authorized editor of Schulich Law Scholars. For more information, please contact hannah.steeves@dal.ca.

AUTONOMOUS WEAPONS SYSTEMS: TAKING THE HUMAN OUT OF INTERNATIONAL HUMANITARIAN LAW

James Foy*

ABSTRACT

Once confined to science fiction, killer robots will soon be a reality. Both the USA and the UK are currently developing weapons systems that may be capable of autonomously targeting and killing enemy combatants within the next 25 years. According to Additional Protocol I to the Geneva Convention and customary international law, weapons systems must be capable of operating within the principles of International Humanitarian Law (IHL). This paper will demonstrate that without significant restrictions on the use of autonomous weapons systems (AWS) or the creation of a new legal framework, the use of AWS is problematic. First, there are legitimate concerns that AWS are, by their nature, incapable of adhering to IHL principles. Second, there is a more fundamental problem: the principles of IHL are actually insufficient to address the unique concerns regarding AWS. Finally, the solutions proposed by proponents of AWS do not sufficiently address these concerns. A legal solution beyond the general principles of IHL must be developed.

Citation: (2014) 23 Dal J Leg Stud 47.

* J.D. 2014, Schulich School of Law, Dalhousie University. The author wishes to thank Professor Robert Currie and Aaron Dewitt.

Introduction

The term Autonomous Weapons Systems (AWS) conjures up images of Terminator style robots: lethal machines with complicated artificial intelligence, capable of killing humans without being hindered by human emotion or cultural constraints. This picture is more science fiction than reality, but current challenges raised by the development of AWS are now at the forefront of international legal discourse and their compliance with International Humanitarian Law (IHL) are real and must be addressed.

In October 2010, a United Nations human rights investigator recommended “[t]he international community urgently...address the legal, political, ethical and moral implications of the development of lethal robotic technologies.”¹ On November 19, 2012, Human Rights Watch, a division of the International Human Rights Clinic, released a report, *Losing Humanity: The Case against Killer Robots*, calling for a ban on the production and use of AWS.² Days later, the US Department of Defence (DoD) released Directive 3000.09, outlining the DoD’s policies on the development and use of AWS.³ On April 9, 2013, the UN special rapporteur on extrajudicial, summary or arbitrary executions called for a moratorium on the development of AWS until a legal framework is developed.⁴ While Human Rights Watch, the DoD and the UN disagree on a solution, they all begin with the presumption that AWS will raise challenges of adherence to IHL.

This paper will demonstrate that the principles of IHL, particularly the principles of distinction and proportionality, are not adequate to address the concerns raised by AWS. Part I will define AWS, distinguish between automatic and autonomous systems, and provide an overview of the current and future use of semi-autonomous and autonomous systems. Part II will outline the principles of distinction and proportionality in IHL. Part III will analyze the challenges of adherence to IHL principles faced by the use of AWS. It will address some, but not all, of the moral objections to AWS.⁵ Part IV will evaluate current proposals for operational solutions and offer legal solutions to ensure that the use of AWS does not violate the principles of IHL. This paper concludes that the principles of IHL are insufficient on their own and that an additional legal framework is necessary to ensure the legal use of AWS.

The analysis in this paper is confined to *lethal* AWS. Non-lethal robots raise their own concerns, particularly in the area of privacy, but they are outside of the scope of this paper. This paper focuses on the legal and moral implications of transferring the decision to kill from human to machine, rather than the wider implications of the automatization of robotic technology. There are also concerns that the existing principles of command responsibility are not sufficient to ensure that AWS comply with the principles of IHL. This topic is beyond of the scope of this paper.⁶

¹ Patrick Worsnip, “UN Official Calls for Study of Ethics, Legality of Unmanned Weapons”, *The Washington Post* (24 October 2010), online: Washington Post <<http://www.washingtonpost.com>>.

² Human Rights Watch, *Losing Humanity: The Case Against Killer Robots*, online: International Human Rights Clinic <http://www.hrw.org/sites/default/files/reports/arms1112ForUpload_0_0.pdf>.

³ US, Department of Defense, *Directive Number 3000.09: Autonomy in Weapon Systems* (21 November 2012), online: <<http://www.dtic.mil/whs/directives/corres/pdf/300009p.pdf>> [DoD 3000.09].

⁴ *Report of the Special Rapporteur on extrajudicial, summary or arbitrary executions*, Christof Heyns, UNHRC, 23d Sess, UN Doc A/HRC/23/47 (2013).

⁵ For a thorough introduction to the challenges raised by AWS, see Armin Krishnan, *Killer Robots: Legality and Ethicality of Autonomous Weapons* (Surrey: Ashgate Publishing Limited, 2009).

⁶ For an introduction to this topic, see Robert Sparrow, “Killer Robots” (2007) 24:1 *Journal of Applied Philosophy* 62.

I. WHAT ARE AUTONOMOUS WEAPONS SYSTEMS?

It is necessary to distinguish between weapons that are automated and weapons that are truly autonomous. The term “autonomous” can be difficult to define. It suggests highly intelligent robots that are capable of individual decision-making. In reality, it looks less like science fiction and more like every day robotics.⁷ AWS may be much closer in operation to a driverless car than to a Terminator.

Defining AWS

Robotist Noel Sharkey defines an **automatic machine** as one that “carries out a pre-programmed sequence of operations or moves in a structured environment.”⁸ By contrast, an **autonomous machine** operates in an unstructured environment. In essence, what makes a machine autonomous is the environment it operates in, rather than its internal processes.

The DoD adopts a broad definition of AWS:

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 input after activation.⁹

The DoD definition’s essential requirement is that, once activated, it can “select and engage targets” without further human input.¹⁰ Human Rights Watch adopts a similar definition: any robot that can select and engage targets without human input, even if there is human oversight, will qualify as a “fully autonomous robot.”¹¹ These definitions capture the primary characteristic of AWS; that is, humans are not necessary in the targeting decision-making process.

The core difference between automatic and autonomous weapons is predictability. An automatic machine is entirely predictable (barring a failure), whereas an autonomous robot can only be predictable as a series of likely outcomes. This distinction is essential in determining whether AWS are capable of adhering to the principles of IHL.

AWS: New Weapons, or New Soldiers?

The rise of AWS has the potential to develop in two different directions: either as an extension of human soldiers or as a replacement for humans in the battlefield.¹² In

⁷ Noel Sharkey, “Automating Warfare: Lessons Learned from the Drones” (2011) 21 *Journal of Law, Information & Science* 140 at 141 [Sharkey, “Automating Warfare”].

⁸ *Ibid.*

⁹ *DoD 3000.09*, *supra* note 3 at 14.

¹⁰ *Ibid.*

¹¹ Human Rights Watch, *supra* note 2 at 2.

¹² Krishnan, *supra* note 5 at 35.

other words, the distinction is between “weapons that augment our soldiers and those that can become soldiers.”¹³

Currently, the dominant view is that robots will be used only to augment and extend our soldiers’ involvement in war.¹⁴ In this context, AWS distance humans from combat. Instead of being a novel category of weapons, AWS are simply the latest technological advancement that began with the bow and arrow. Similarly, the critical responses to the potential introduction of AWS are not novel. Some see any introduction of new weapons as unethical or illegal.¹⁵

However, the idea that AWS will replace our soldiers is gaining traction. AWS are more than an extension of human combat when they can make decisions to kill without human involvement.¹⁶ While the use of drones may be criticized for other reasons, their capability of adhering to the principles of IHL is uncontroversial because humans are involved in the targeting process. AWS would take human operators out of the decision-making loop.¹⁷ Distancing humans from war through technology has been a common theme of weapons development, but taking humans out of the loop completely is a fundamental shift in the development of weapons systems.

The Future of AWS

Although AWS are not yet a reality on the battlefield, the level of autonomy in weapons has been growing steadily and there are several weapons systems approaching fully autonomous capabilities.¹⁸ Experts believe that their introduction is inevitable and imminent.¹⁹ The former chief scientist of the US Air Force contends that the technology required for “fully autonomous military strikes” already exists.²⁰

The development of AWS will take place incrementally, beginning with aspects of operations such as take off and navigation, leading to full autonomy over time.²¹ As technological advances are made, increasingly sophisticated sensing and computational systems will be implemented. The increased tempo of warfare and pressures to minimize a state’s own military casualties will also create a demand for AWS.²²

Several weapons systems already include semi-autonomous capabilities and the level of automation in weapons systems is steadily increasing. The South Korean military recently deployed a stationary sentry robot in the Korean Demilitarized Zone that is capable of detecting and selecting targets. It can respond with lethal or non-lethal force, depending on the circumstances. According to the developer, “the ultimate decision

¹³ Major David F Bigelow, “Fast forward to the robot dilemma” *Armed Forces Journal* (1 November 2007), online: <<http://www.armedforcesjournal.com>>.

¹⁴ Krishnan, *supra* note 5 at 35.

¹⁵ Kenneth Anderson & Matthew Waxman, “Law and Ethics for Robot Soldiers” (2012) 176 Policy Review 35 at 39-40.

¹⁶ Krishnan, *supra* note 5 at 33.

¹⁷ Markus Wagner, “Taking Humans Out of the Loop: Implications for International Humanitarian Law” (2011) 21 Journal of Law, Information & Science 155 at 157-8.

¹⁸ See Human Rights Watch, *supra* note 2. See also Timothy Coughlin “The Future of Robotic Weaponry and the Law of Armed Conflict: Irreconcilable Differences?” (2011) 17 UCL Jurisprudence Review 67.

¹⁹ See especially Gary E Marchant et al, “International Governance of Autonomous Military Robots”, online: (2011) 12 Colum Sci & Tech L Rev 272 <<http://www.stlr.org>>.

²⁰ Werner JA Dahm, “Killer Drones Are Science Fiction”, *The Wall Street Journal* (15 February 2012) 11.

²¹ Sharkey, “Automating Warfare”, *supra* note 7 at 141.

²² Anderson & Waxman, *supra* note 15 at 36.

about shooting should be made by a human, not the robot.”²³ However, the robot is capable of making that decision without human input.²⁴

The Phalanx Close In Weapons Systems for *Aegis* class cruisers in the US Navy is currently “capable of autonomously performing its own search, detect, evaluation, track and kill assessment functions.”²⁵ The system has four modes: semi-automatic, where humans control the firing decision; automatic special, where humans set targets but the software determines how to carry them out; automatic, where humans monitor the system but it works without their input; and casualty, where the system does whatever is necessary to save the ship.²⁶

The United Kingdom is currently testing a new semi-autonomous aircraft, *Taranis*. The designer, BAE Systems, describes it as “an autonomous and stealthy unmanned aircraft.”²⁷ Although humans will remain in the loop for the time being, it may be capable of autonomous flight.²⁸ The United States of America (US) is also developing a semi-autonomous drone, the *X-47B*, which will be able to take off and land without human input. The developer contends that it is a system that “takes off, flies a pre-programmed mission, and then returns to base in response to mouse clicks from its mission operator. The mission operator monitors the...vehicle’s operation, but does not actively ‘fly’ it via remote control as is the case for other unmanned systems currently in operation.”²⁹ The current development of *X-47B* does not envision autonomous target selection.

The development of AWS has been included in all roadmaps of the US forces since 2004.³⁰ The US Air Force’s Flight Plan suggests that fully autonomous flight systems will be possible as early as 2025.³¹ Sharkey claims to have read valid robotics development reports from over 50 countries that are currently developing autonomous weapons systems, including Canada.³² US Air Force Major Michael A. Guetlin states that “[it] is not a matter of ‘will’ we employ [autonomous weapons]; it is a matter of ‘when’ we employ them.”³³

²³ Jean Kumagai, “A Robotic Sentry for Korea’s Demilitarized Zone” *IEEE Spectrum* (1 March 2007), online: IEEE Spectrum <<http://www.spectrum.ieee.org>>.

²⁴ Although the robot is capable of selecting and engaging targets without human input, its location in the DMZ makes it unnecessary for the robot to distinguish between civilian and enemy combatant. Any person that crosses a pre-determined line is considered an enemy combatant by the robot.

²⁵ United States Navy, *Fact File: MK 15 – Phalanx Close-In Weapons System (CIWS)*, (19 Oct 2012), online: <<http://www.navy.mil>>.

²⁶ Marchant, *supra* note 19 at 287.

²⁷ BAE Systems, *Taranis*, online: BAE Systems <<http://www.baesystems.com>>.

²⁸ Human Rights Watch, *supra* note 2 at 17-18.

²⁹ Northrop Grumman, “Unmanned Combat Air System Carrier Demonstration” at 2, online: Northrop Grumman Corporation <http://www.northropgrumman.com/Capabilities/X47BUCAS/Documents/X-47B_Navy_UCAS_FactSheet.pdf>.

³⁰ Noel Sharkey, “The Automation and Proliferation of Military Drones and the Protection of Civilians” (2011) 3(2) *Law, Innovation and Technology* 229 at 235 [Sharkey, “Automation and Proliferation”].

³¹ United States Air Force, *Unmanned Aircraft Systems Flight Plan 2009-2047* (18 May 2009) at 50, online: <http://www.fas.org/irp/program/collect/uas_2009.pdf>.

³² Sharkey, “Automation and Proliferation”, *supra* note 30 at 231.

³³ Major Michael A Guetlin, *Lethal Autonomous Weapons – Ethical and Doctrinal Implications*, (JMO Department, Naval War College, 2005) [unpublished] at 18, online: Defence Technical Information Center <<http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA464896>>.

The Perceived Benefits of AWS

Gordon Johnson, a member of the now-defunct Pentagon Joint Forces Command emphasized the benefits of AWS: “They don’t get hungry. They’re not afraid. They don’t forget orders. They don’t care if the guy next to them has just been shot. Will they do a better job than humans? Yes.”³⁴

There are a number of tactical and operational factors that promote the development of lethal AWS.³⁵ AWS are cheaper to operate than human operated weapons and are capable of operating continuously, without the need for rest.³⁶ Although it can be possible to extend mission times for humans up to 72 hours with performance enhancers, eventually a human needs rest.³⁷ AWS are capable of long-term performance as their batteries sustain them. As battery and recharging technology improves, the possible mission time for AWS will continue to grow.

Fewer humans are needed for the operation of AWS.³⁸ It may soon be possible for a single operator to manage a swarm of semi-autonomous drones or for a single human commander to assign mission parameters to AWS and monitor them from a safe distance. This distances the human warfighter and expands the battle space. It will be possible to conduct combat over a much larger area than before.

AWS are also potentially capable of processing battlefield information faster and more efficiently than humans.³⁹ AWS can be fitted with any variety of sensory technologies, including: infrared vision, sonar, high definition cameras and sophisticated auditory sensors. This would give AWS an advantage over human sensory capabilities.

One weakness of current remotely piloted vehicles is the possibility that the enemy will interfere with their satellite or radio links.⁴⁰ AWS alleviates this concern, as they will be capable of operating without continuous contact with home base. Remotely piloted systems currently have a delay time of approximately 1.5 seconds, limiting their effectiveness in a higher tempo battle space. This delay would make it impossible for a remotely piloted system to engage in an aerial dogfight, while autonomous flight capabilities would make this possible.

Proponents of AWS suggest that AWS may in fact be more capable of adhering to the principles of IHL than human soldiers.⁴¹ They may be able to act more conservatively because they will not have a need for self-preservation. Robotic sensors will be better equipped to make battlefield observations than humans. AWS lack the emotions that can cloud a human’s judgment. They will be immune to the psychological problem of **scenario fulfillment**, the phenomenon of humans using new information to fit a pre-existing belief pattern.

The introduction of AWS into the battlefield is inevitable, but it will be incremental. Although humans will be in the loop as a fail-safe when AWS are first deployed, their involvement will diminish over time. As human involvement diminishes, the diffi-

³⁴ Tim Weiner, “‘GI Robot’ Rolls Toward the Battlefield” *New York Times* (1 February 2005), online: New York Times <<http://www.nytimes.com>>.

³⁵ Marchant, *supra* note 19 at 275.

³⁶ Guetlin, *supra* note 33 at 2.

³⁷ Krishnan, *supra* note 5 at 41.

³⁸ Marchant, *supra* note 19 at 275.

³⁹ Guetlin, *supra* note 33 at 4-5.

⁴⁰ Noel Sharkey, “Saying No! to Lethal Autonomous Targeting” (2010) 9:4 *Journal of Military Ethics* 369 at 377 [Sharkey, “Saying No”].

⁴¹ See especially Marchant, *supra* note 19 at 279-280.

culties faced by AWS in adhering to the principles of IHL will become more and more significant, requiring a thorough legal analysis.

II. THE PRINCIPLES OF INTERNATIONAL HUMANITARIAN LAW (THE LAW OF ARMED CONFLICT)

Warfare is governed by IHL, also known as the Law of Armed Conflict.⁴² IHL is relevant to the legality of weapons in two ways. First, a weapon may be incapable of adhering to the principles of IHL, rendering it illegal *per se*; even when it is deployed against a lawful target, the weapon will be illegal. Second, the weapon can be used in a way that is unlawful.⁴³ For example, a rifle is a lawful weapon, but its use is unlawful if used to shoot civilians.

Some analyses of AWS have conflated the two methods by which they could potentially be rendered illegal. Human Rights Watch's paper, *Losing Humanity*, does not mention this distinction and has been criticized for oversimplifying the application of IHL.⁴⁴ It is important to be clear about how AWS may violate the principles of IHL so that concerns can be adequately addressed. It is equally important to avoid overemphasizing the failure to address the differences between illegality *per se* and illegality by use. Doing so risks undervaluing what is really at issue: the possibility that AWS will be incapable of adhering to cardinal principles of IHL in some, if not all, circumstances.

Illegal *per se*

A weapon will be illegal *per se* if it causes superfluous injury or unnecessary suffering or is wholly incapable of adhering to the principles of IHL.⁴⁵ This restriction is very limited because most weapons will be capable of adhering to IHL principles in specific circumstances.

For example, the International Court of Justice (ICJ) in an advisory opinion considered whether the use of nuclear weapons was illegal *per se*.⁴⁶ The court concluded that nuclear weapons were not inherently incapable of distinction or proportionality, nor would they cause superfluous injury or unnecessary suffering in all circumstances, and so were not illegal *per se*.⁴⁷ If nuclear weapons, which are among the most deadly, are not illegal *per se*, then it is highly unlikely that AWS would be.

⁴² Gary D Solis, *The Law of Armed Conflict: International Humanitarian Law in War* (New York: Cambridge University Press, 2012) at 2.

⁴³ Michael N Schmitt, "Autonomous Weapons Systems and International Humanitarian Law: A Reply to Critics" *Harvard National Security Journal Features* (5 February 2013) at 3 online: Harvard National Security Journal <<http://harvardnsj.org/wp-content/uploads/2013/02/Schmitt-Autonomous-Weapon-Systems-and-IHL-Final.pdf>>.

⁴⁴ *Ibid* at 8.

⁴⁵ See William Boothby, *Weapons and the Law of Armed Conflict*, (New York: Oxford University Press, 2009) at ch 5.

⁴⁶ *Legality of the Threat or Use of Nuclear Weapons*, Advisory Opinion [1996] ICJ Rep 226, [*Nuclear Weapons Advisory Opinion*].

⁴⁷ Meredith Hagger & Tim McCormack, "Regulating the Use of Unmanned Combat Vehicles: Are General Principles of IHL Sufficient?" (2011) 21 *Journal of Law, Information and Science* 74 at 81-84.

Illegal by use

This paper will focus on the second method by which a weapons system could be illegal: if the principles of IHL are violated through its use. It is important to recognize that if AWS violated the principles of IHL in certain situations, only their use in those situations would be illegal, not the use of AWS altogether.

There are four core IHL principles that apply to every combat operation: distinction, proportionality, military necessity and unnecessary suffering.⁴⁸ This paper will only address distinction and proportionality. While the use of AWS engages the principles of military necessity and unnecessary suffering, these principles are engaged in a different way than distinction and proportionality. The difficulties of adherence to the principles of distinction and proportionality naturally arise when humans are removed from the decision-making loop. The principles of military necessity and unnecessary suffering are less affected by the removal of a human from the loop and more situation-specific.

The Principle of Distinction

There are two components to the principle of distinction (sometimes referred to as discrimination): combatants must be able to distinguish (i) between civilians and enemy combatants, and (ii) between civilian and military objects.⁴⁹ This principle is codified in Article 48 of the Additional Protocol I to the Geneva Convention:

In order to ensure respect for and protection of the civilian population and civilian objects, the Parties to the 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.⁵⁰

For states that are not signatories to Additional Protocol I, the principle applies as customary international law.⁵¹ The *Commentary on the Additional Protocols*, produced by the International Committee of the Red Cross (ICRC), holds that Article 48 of Additional Protocol I reflects the foundational principle of the laws and customs of war that civilians must be protected and therefore must be distinguished from combatants.⁵² In addition, in the *Nuclear Weapons Advisory Opinion*, the ICJ held that the rule against indiscriminate attacks is a “cardinal” principle of IHL.⁵³

The prohibition on indiscriminate attacks is particularly concerned with the attacker’s doubt. Where there is sufficient doubt, a target will be presumptively immune from attack.⁵⁴ This is codified in Article 50(1) of Additional Protocol I⁵⁵ and has also been

⁴⁸ Solis, *supra* note 42 at 250.

⁴⁹ *Ibid* at 251.

⁵⁰ *Protocol Additional to the Geneva Convention of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts (Protocol I)*, 8 June 1977, 1125 UNTS 3, art 48, (entered into force 7 December 1979) [*Additional Protocol I*].

⁵¹ Solis, *supra* note 42 at 251.

⁵² Yves Sandoz, Christophe Swinarski & Bruno Zimmermann, eds, ICRC, *Commentary on the Additional Protocols of 8 June 1977 to the Geneva Conventions of 12 August 1949* (Geneva: Martinus Nijhoff Publishers, 1987), at 598.

⁵³ *Nuclear Weapons Advisory Opinion*, *supra* note 46 at 78.

⁵⁴ Schmitt, *supra* note 43 at 16.

⁵⁵ *Additional Protocol I*, *supra* note 50, art 50(1).

accepted as customary international law.⁵⁶ Although some level of doubt is permissible, the presumption will be created in situations that cause “a reasonable attacker in the same or similar circumstances to hesitate before attacking,”⁵⁷

Adherence to the principle of distinction has become increasingly difficult, as military operations have evolved from state against state warfare to counterinsurgency operations.⁵⁸ However, the challenges of applying the principle of distinction do not change the core of the principle; namely, parties to a conflict must distinguish between civilian targets and military targets.

The Principle of Proportionality

The principles of IHL strive to protect civilian populations, but there is no way to eliminate civilian death and injury from war altogether. Proportionality seeks to address the protection of civilians directly and mandates that where collateral damage to civilians occurs, it must be proportional to military advantage.⁵⁹ The rule of proportionality is defined in Article 51(5)(b) of Additional Protocol I. It states that a violation of proportionality will be “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 advantage anticipated.”⁶⁰ The ICRC’s study of customary IHL restated the principle in these terms:

[The] armed forces and their installations are objectives that may be attacked wherever they are, except when the attack could incidentally result in loss of human life among the civilian population, injuries to civilians, and damage to civilian objects which would be excessive in relation to the expected direct and specific military advantage.⁶¹

There is no reference to proportionality in Additional Protocol II, which applies to intranational armed conflicts.⁶² However, the ICRC argues that because proportionality is inherent to the principle of humanity, which is included in the Protocol’s preamble, it must be included in the Protocol’s application. In addition, the ICRC could find no official practice contrary to the principle of proportionality in either international or intranational armed conflicts, arguing that the principle has crystallized into customary law.⁶³

“Military advantage” has been interpreted by many states, including Canada and the US, to include the particular advantage anticipated from an attack, as well as the advantage anticipated to the military operation as a whole.⁶⁴ The military advantage

⁵⁶ See Sandoz, Swinarski & Zimmerman, *supra* note 52, at 1410-1439.

⁵⁷ Schmitt, *supra* note 43 at 16.

⁵⁸ Solis, *supra* note 42 at 254.

⁵⁹ *Ibid* at 274.

⁶⁰ *Additional Protocol I*, *supra* note 50, art 51(5)(b).

⁶¹ Sandoz, Swinarski & Zimmerman, *supra* note 52 at 620-1.

⁶² *Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of Non-International Armed Conflicts (Protocol II)*, 8 June 1977, 1125 UNTS 609 (entered into force 7 December 1978).

⁶³ Jean-Marie Henckaerts et al, eds, *Customary International Humanitarian Law*, vol 1, ICRC (Cambridge: Cambridge University Press, 2005) at 48.

⁶⁴ *Ibid* at 49.

must be “concrete and direct.” According to the *Commentary on the Additional Protocols*, the phrase “concrete and direct” indicates that the advantage “must be substantial and relatively close, and that advantages which are hardly perceptible and those which would only appear in the long term should be disregarded.”⁶⁵ Canada’s *Law of Armed Conflict Manual* states that a “concrete and direct” advantage will exist where “the commander has an honest and reasonable expectation that the attack will make a relevant contribution to the success of the overall operation.”⁶⁶

Proportionality requires a contextual weighing of two factors: the possibility of harm to civilians and civilian objects and the potential military advantage of the attack. Potential harm to civilians is more readily capable of objective determination. Commanders already use collateral damage simulators to ensure attacks are proportional.⁶⁷ The determination of military advantage, on the other hand, is more contextual and discretionary. In determining whether the military advantage requirement has been met, one asks if a ‘reasonable commander’ would arrive at a similar conclusion. The evaluation of military advantage on the basis of the reasonable commander allows for operational discretion. The contextual and discretionary nature of proportionality is what causes concerns that AWS may be incapable of adhering to the principle.

Although distinction and proportionality are distinct concepts of IHL, they are intertwined. Distinction requires combatants to distinguish between enemy combatants and civilians. So, a combatant cannot intend to harm civilians, but proportionality enables them to attack knowing that some civilians will be harmed.⁶⁸ This is important to keep in mind when considering the challenges AWS faces in complying with IHL. If AWS are incapable of distinction, they will also be incapable of proportionality.

III. CHALLENGES OF COMPLIANCE WITH IHL

Generally, technological advances in warfare have outpaced the development of IHL. This phenomenon is not unique to IHL, but common wherever legal regimes interact with technological advancements.⁶⁹ For AWS to be used legally, they must be capable of adhering to the principles of IHL, including distinction and proportionality.

Compliance with the Principle of Distinction

On its face, the principle of distinction can be seen as an objective requirement. AWS must be able to objectively assess whether a potential target is a civilian target or a military target.⁷⁰ The distinction looks like a black and white rule; either a target is or is not a military target. However, difficulties arise because a target can be classified as both a civilian and military target depending on the context.

⁶⁵ Sandoz, Swinarski & Zimmerman, *supra* note 52 at 2209.

⁶⁶ Canada, Office of the Judge Advocate General, *The Law of Armed Conflict at the Operational and Tactical Level*, 2001, s 415(2), online: Office of the Judge Advocate General <<http://www.forces.gc.ca/jag>>.

⁶⁷ Schmitt, *supra* note 43 at 19-20.

⁶⁸ Solis, *supra* note 42 at 276.

⁶⁹ Wagner, *supra* note 17 at 157.

⁷⁰ *Ibid* at 159.

In *Killer Robots*, Armin Krishnan identifies three main concerns regarding the ability of AWS to distinguish legal targets from civilian targets: (i) AWS may be susceptible to “weak machine perception;” (ii) AWS may have difficulties in interacting with their environment, leading to the “frame problem;” and (iii) there may be a problem of “weak software.”⁷¹

Weak Machine Perception

Distinction requires an evaluation based on sensory input. Current technology is only beginning to approach the ability to distinguish between human and non-human objects, never mind between civilians and combatants.⁷² This suggests that while it may be technically possible for AWS to be capable of distinction, it will take some time before the capability to distinguish is a reality.

This problem is compounded in intranational armed conflicts with non-uniformed enemy combatants. In such situations, a target is only lawful if it is directly engaged in hostile activity or intends to engage in hostile activity. An AWS targeting decision would have to be based on situational awareness and an understanding of human intention.⁷³ Non-uniformed (and consequently unlawful) combatants in an armed conflict are identified by their engagement or intention to engage in hostilities.⁷⁴ One solution to this problem would be to only allow AWS to fire when they have been fired upon. Another would be to limit the use of AWS to situations where the declared hostile force is easily recognizable.⁷⁵ A third way to ensure AWS adhere to the principle of distinction would be to limit the number of potential targets to a fixed list of lawful targets.⁷⁶

Even if AWS are wholly incapable of distinguishing between civilians and combatants, it would be possible to use them against these lawful targets in a battle space that does not contain civilians or civilian objects.⁷⁷ This approach is technically correct, but may be unrealistic given the trend of warfare towards counterinsurgency operations in or near civilian spaces. As a result, limiting the use of AWS to battle spaces void of civilians would render AWS unfit for use in almost all circumstances. Given the perceived benefits of AWS, it is unlikely that this limitation would be adopted by any state.

John Canning, a Combat Systems Engineer for the Unmanned Systems Integration Branch at the US Naval Surface Warfare Center, has proposed a possible solution to some of the issues raised by AWS. In his proposal, AWS would target the weapon, rather than the human holding the weapon, so any injury to the human would be considered collateral damage.⁷⁸ Although theoretically possible, this solution does not adequately address the possibility of indiscriminate attacks. Distinguishing between a

⁷¹ Krishnan, *supra* note 5 at 98-99.

⁷² Sharkey, “Automating Warfare”, *supra* note 7 at 143-144.

⁷³ Sharkey, “Saying No”, *supra* note 40 at 379.

⁷⁴ Solis, *supra* note 42 at 254-255. Combatants are under a duty to distinguish themselves from civilians. Art 44.3 of Additional Protocol I requires it. Terrorist organizations do not typically distinguish themselves, making distinction in modern warfare more difficult than it has been in traditional warfare.

⁷⁵ Major Jeffrey S Thurnher, “No One at the Controls: Legal Implications of Fully Autonomous Targeting” (2012) 67:4 Joint Forces Quarterly 77 at 83.

⁷⁶ Anderson & Waxman, *supra* note 15 at 41.

⁷⁷ Schmitt, *supra* note 43 at 11.

⁷⁸ John S Canning, “A Concept of Operations for Armed Autonomous Systems” (2006), online: Defense Technical Information Center <http://www.dtic.mil/ndia/2006disruptive_tech/canning.pdf>.

weapon and any other object may be just as difficult as distinguishing between a civilian and an enemy combatant.⁷⁹ The potential problems with this solution are illustrated with a simple example. If enemy combatants force children or other civilians into transporting weapons for them, they would not be legal targets. However, under Canning's proposal, they would be treated as collateral damage.

Frame Problem

In a complex and fast-paced modern battle space, AWS will have difficulty interpreting all the information needed to correctly assess the situation. Processing all of the possible scenarios would take an excessive amount of time. Consequently, AWS will have to be programmed to distinguish between relevant and irrelevant information. In an open environment, programming this type of distinction could lead to situations where the information is incorrectly interpreted, causing an indiscriminate attack. According to Armin Krishnan, this means AWS would be too slow to be militarily effective, or else would be prone to indiscriminate use because the systems "would often miss important details or incorrectly interpret situations."⁸⁰

The frame problem is complicated by the rule that an attack will be unlawful where the legitimacy of the target is in significant doubt. Any doubt about the legitimacy of a target does not create a presumption that the target is unlawful: instead, the doubt must cause "a reasonable attacker in the same or similar circumstances to hesitate before attacking."⁸¹ The threshold is framed in terms of human reasonableness, which complicates its adoption in AWS. This determination is contextual and would require different "doubt thresholds" depending on the circumstances. Michael Schmitt suggests that as long as human operators do not program the "doubt thresholds" unreasonably high (so that the AWS is more likely to attack), AWS will not violate the principles of distinction.⁸²

Weak Software

As software becomes more complicated, it becomes less predictable. No one programmer understands or knows the entire piece of software, so interactions within it become unpredictable as well.⁸³ Combined with an open environment, this could lead to situations where AWS apply force indiscriminately because of an unanticipated software error.

In 1997, Murray Campbell and Feng-hsiung Hsu created *Deep Blue*, a chess playing computer that eventually beat top rated chess player Garry Kasparov. The difficulties faced by Campbell and Hsu while programming *Deep Blue* exemplify the weak software problems faced by programming AWS. At a certain point, the computer became a more capable chess player than they were, making it increasingly difficult to tell if a move was

⁷⁹ Krishnan, *supra* note 5 at 106.

⁸⁰ *Ibid* at 99.

⁸¹ Schmitt, *supra* note 43 at 16.

⁸² *Ibid* at 17.

⁸³ Krishnan, *supra* note 5 at 100.

a bug or good tactics.⁸⁴ The same difficulty will be faced when programming AWS: at a certain point it will become difficult to tell if the machine is making an error or if it is seeing something that a human cannot.

Sharkey argues that AWS will not be able to discriminate between combatants and civilians. Although it is technically possible to program AWS to avoid civilian targets, this is only achievable if there is a clear definition of “civilian.”⁸⁵ In non-international armed conflict this definition becomes less and less clear. Imagine a situation where terrorists are forcing occupants of a village to transport weapons. The villagers are carrying weapons and could be considered by AWS as participating in a hostile act as enemy combatants. However, more subjective factors, like body language that indicates the villagers are transporting the weapons against their will, may be missed by AWS. Consequently, AWS may have difficulty in correctly assessing the situation and avoiding unnecessary civilian death.

Although the principle of distinction appears to be an objective requirement, its subjective elements create challenges for the use of AWS. First, the sensory technology must develop sufficiently so that AWS have enough information to be capable of distinguishing between civilian and military targets. Second, that information must be processed efficiently and accurately, so that mistakes are not made and the AWS target indiscriminately. Both of these steps require technology that does not yet exist. It is conceivable that the technology required to distinguish will be developed, but the challenges raised by removing humans from the targeting decision cannot be ignored simply because the technology may exist someday.

Compliance with the Principle of Proportionality

It is difficult to establish black and white rules with respect to proportionality.⁸⁶ A report described this problem to the International Criminal Tribunal for the Former Yugoslavia: “[one] cannot easily assess the value of innocent human lives as opposed to capturing a military objective.”⁸⁷ Adherence to the principle of proportionality also requires a subjective assessment. It is difficult to apply in practice and requires a weighing of potentially competing interests: military advantage and the protection of civilians.⁸⁸ This weighing of interests is only possible on a case-by-case basis: different circumstances require different responses.

The evaluation of proportionality requires relative weight be placed on competing interests. In order to analyze a situation and deliver a proportional response, AWS face several challenges. They must be able to anticipate the effect of all potential decisions and how many civilian casualties could result. They would also have to react to changing circumstances.⁸⁹ Then they must calculate the military advantage and determine whether the collateral damage is acceptable.

⁸⁴ Nate Silver, *The Signal and the Noise: Why Most Predictions Fail – but Some Don't* (New York: The Penguin Press, 2012) at 286.

⁸⁵ Sharkey, “Automating Warfare”, *supra* note 7 at 143-144.

⁸⁶ Solis, *supra* note 42 at 274.

⁸⁷ *Final Report to the Prosecutor by the Committee Established to Review the NATO Bombing Campaign Against the Federal Republic of Yugoslavia*, (8 June 2000) at para 48 (International Criminal Tribunal for the Former Yugoslavia), online: ICTY <<http://www.icty.org>>.

⁸⁸ See Wagner, *supra* note 17 at 159, 162.

⁸⁹ *Ibid* at 163.

According to William Boothby, the proportionality rule has no direct application to weapons development because of the requirement for case-by-case determination.⁹⁰ However, unlike other weapons, AWS could replace the human decision maker. While humans may be capable of balancing complex interests, the same cannot be said for AWS. Therefore, proportionality must be considered in questioning the legality of AWS.

Collateral Damage

Systems that determine the likelihood of collateral damage already exist and are used to determine what level of command is required to authorize an attack. A commander weighs the potential collateral damage against military advantage.⁹¹ The same frame problem exists when AWS operate in open and unstructured environments. To calculate the collateral damage of an attack, AWS will either have to calculate the consequences of every possible action (taking an excessive amount of time) or make assumptions that could potentially lead to a disproportionate attack. Determinations of collateral damage will always involve assumptions; certainty is almost never possible in armed conflict. If AWS are employed in open civilian environments, the information relied upon to support assumptions in a collateral damage determination must be collected and processed adequately.

Military Advantage

Currently, no system is capable of calculating military advantage, but proponents suggest that it is theoretically possible.⁹² The frame problem complicates any determination of military advantage because the decision maker would need to consider the immediate and long-term consequences of an action. This ability has yet to be replicated by software. Since AWS do not have an infinite amount of time to make these calculations, some shortcuts will have to be programmed into the software, potentially leading to errors and disproportionate attacks.

Military advantage and collateral damage are constantly shifting and depend on the context. An example illustrates the potential challenges faced by AWS: if an enemy combatant is setting up a defensive position on top of a building, there will be a military advantage in targeting that combatant. If there are no civilians in the area, the probability of collateral damage will be sufficiently low and the attack will be proportional. However, if a large group of civilians runs into the building, the potential for collateral damage becomes unacceptable and the attack will not be proportional. What would be immediately obvious to a human soldier requires complex processing and sensing capabilities, as well as an algorithm that is capable of making speedy and correct determinations of proportionality.

⁹⁰ Boothby, *supra* note 45 at 79.

⁹¹ Schmitt, *supra* note 43 at 19-20.

⁹² *Ibid.*

Balancing the Two: The Reasonable Commander

Marcus Wagner has suggested the challenges faced in programming AWS may render their use “almost useless except in the narrowest of circumstances.”⁹³ If they cannot be programmed to meet the reasonable commander requirement (i.e. to balance the potential for collateral damage with a calculation of military advantage), then they will never be capable of a proportional attack.

Weighing collateral damage and military advantage requires the evaluation of a multitude of factors. A complete understanding of the risks associated with AWS may be impossible. The balancing of multiple factors would involve complex programming and it may not be possible to predict outcomes with any certainty.⁹⁴ Complex software is written not by one programmer, but by hundreds. Unforeseen interactions of code may result in undesirable results, especially because AWS will be deployed in open and unstructured environments.

Major A. Guetlin suggests that adhering to the principle of proportionality is really a question of probabilities: “If the probability of success is low, or the probability of excessive collateral damage is high, then the weapon system will not engage.”⁹⁵ If AWS are operating in a civilian centre, the commander must set the threshold for engagement higher than if they were operating in a desert. Provided that the commander has programmed the AWS correctly, Major A. Guetlin argues that their use would be proportional.⁹⁶ This would put control of proportionality back in the hands of a human. However, it ignores the challenges of predetermining collateral damage and the probabilities of success accurately in advance of the mission and sidesteps the issue. While AWS operating under this probabilities approach may end up being more proportional, the underlying potential for disproportionality is not addressed.

There is also a fundamental moral objection to AWS: taking the decision to kill away from a human and giving it to machines. Even a flawed human being is more capable of moral action than a robot without a conscience.⁹⁷ AWS would have no awareness beyond their own internal processes and would have no concept of the finality of life.⁹⁸ Due to this inherent limitation, AWS would be incapable of acting proportionally. Even if collateral damage and military advantage are capable of numerical calculation, if AWS cannot comprehend the human consequences of its actions beyond numbers on a balance sheet, they will not be capable of meaningful compliance with the principle of proportionality.

It is conceivable that AWS will one day be capable of distinction and proportionality in some circumstances. Proponents of AWS have argued that by limiting the battle spaces AWS participate in or their potential targets, compliance with the principle of distinction is possible. By programming doubt thresholds into AWS, they may also be capable of proportionality. However, the proposed operational solutions avoid addressing the legitimate challenges faced by the use of AWS in modern warfare by placing potentially unrealistic restrictions on their use. One only needs to look at the prevalence of unmanned drones to imagine the potential growth of AWS usage in war.

⁹³ Wagner, *supra* note 17 at 163.

⁹⁴ Marchant, *supra* note 19 at 284.

⁹⁵ Guetlin, *supra* note 33 at 11.

⁹⁶ *Ibid* at 12.

⁹⁷ Anderson & Waxman, *supra* note 15 at 42.

⁹⁸ Krishnan, *supra* note 5 at 132-133.

Additional Challenges

The principles of distinction and proportionality do not operate in watertight compartments. There are additional challenges raised by the implementation of AWS that engage both principles: empirical skepticism in the value of AWS, the potential expansion of the battle space, the risk of moral disengagement, and the concerns of damaging civilian relations.

Empirical Skepticism

Robotic technology may never reach the point of being able to adhere to the principles of distinction or proportionality. The promise of ever-increasing capabilities of robotics that will “overcome human failings” is a slippery slope and may lead to the introduction of AWS before sufficient safeguards are in place.⁹⁹ It would be unwise to count AWS out altogether; it is impossible to predict how technology will advance in the next 30 years. However, the possibility that adequate technology may never be developed must be considered.

Expansion of the Battle Space

If there is no risk to military personnel, then the human cost of going to war will be significantly lowered. This could lead to the expansion of participation in armed conflict.¹⁰⁰ This objection is not unique to AWS; it has also been raised by critics of remotely piloted drones.¹⁰¹ However, remotely piloted drones require constant communication with their home base, whereas AWS can operate independently. Therefore, AWS may have an even larger impact on the expansion of military intervention than drones.

Moral Disengagement

The use of AWS may lead to moral disengagement due to the distancing of humans from battle. AWS mitigate two major obstacles faced by soldiers: “fear of being killed and resistance to killing.”¹⁰² Peter Singer interviewed pilots of unmanned aerial vehicles, more commonly known as drones, for his book, *Wired for War*. One unnamed pilot reportedly said: “The truth is, it isn’t all I thought it was cracked up to be. I mean, I thought killing somebody would be this life-changing experience. And then I did it, and I was like ‘All right, whatever...’”¹⁰³ This young pilot’s experience with distanced killing exemplifies the problem of moral disengagement. When the human overseer does not even have control over the targeting decision, the moral disengagement will only deepen.

⁹⁹ Anderson & Waxman, *supra* note 15 at 42.

¹⁰⁰ See Sharkey, “Automating Warfare”, *supra* note 7.

¹⁰¹ Anderson & Waxman, *supra* note 15 at 44.

¹⁰² Automating Warfare, *supra* note 6 at 145.

¹⁰³ Peter Singer, *Wired for War: The Robotics Revolution and Conflict in the Twenty-First Century* (New York: The Penguin Press 2009) at 391-392.

Civilian Relations

If it is possible for AWS to distinguish between civilians and combatants more capably than a human, then humans may be perceived as less capable. This can lead to distrust and anxiety amongst civilian populations, which could endanger the long-term success of the mission.¹⁰⁴ It is also likely that civilian populations will perceive AWS as being less proportional than humans, compounding the problem.

The potential use of AWS raises challenges of compliance with the principles of distinction and proportionality. The technology may never be capable of meeting standards that are designed for humans. Distinguishing between civilian and enemy combatants requires computational processing power that has not yet been achieved. Even if the technology is developed, the definition of civilian may not be sufficiently precise for the purposes of software. The subjective and contextual nature of proportionality requires a weighing of factors that may not be possible for a machine. In addition, there is a moral objection to AWS: the decision to kill a human should not be given to a robot.

It has been argued that, by evaluating the challenges of strictly complying with IHL principles, AWS are held to a higher standard than humans.¹⁰⁵ History is full of examples of indiscriminate and disproportionate attacks committed by human soldiers. However, this argument is deficient for two reasons. First, the concerns raised in this paper demonstrate that AWS have a long way to go before they are capable of approaching human-like decision-making. If AWS are not required to be better than a human at distinguishing and calculating proportionality, they must at least reach the same level. Second, it ignores the perceived benefits of the introduction of AWS relied upon by proponents of the systems. If AWS are developed in part as a humane development in the exercise of war, they should be held to a higher standard than humans, as they must be more precise and proportional. Only then is the laudable goal of reducing the destruction of war while maintaining military advantage achievable.

IV. POTENTIAL LEGAL RESPONSES TO AWS

Historical reality suggests that if a new weapon has substantial advantages for one state, it will be gradually adopted by other states over time.¹⁰⁶ Given the potential advantages of AWS, it is likely that their use will proliferate in the next 30 years. Although there are serious concerns that must be addressed regarding the adherence to distinction and proportionality principles, it is possible that they could be overcome.

In *Killer Robots*, Armin Krishnan argues that the development of AWS can lead to either the humanizing of war or to a more dangerous battle environment. What path AWS follow “will largely depend on an effective regulation...”¹⁰⁷ If technology develops to the point where AWS are capable of distinction and proportionality in some

¹⁰⁴ *Ibid* at 285.

¹⁰⁵ AWS and IHL, *supra* note 43 at 12.

¹⁰⁶ Anderson & Waxman, *supra* note 15 at 40.

¹⁰⁷ Krishnan, *supra* note 5 at 4.

circumstances, they will not be illegal *per se*. However, they may be illegal in their use. How should the concerns surrounding AWS be addressed so that they will only be used in compliance with the principles of IHL?

Are the Principles of IHL Sufficient?

Michael Schmitt, chairman of the International Law Department of the US Naval War College, suggests that if AWS are not capable of meeting the legal standards of IHL, then they would already be unlawful.¹⁰⁸ However, the mere applicability of IHL principles and attendant criminal liability for breaches thereof may not be sufficient to ensure that AWS are used lawfully.

This fear is not limited to AWS. Many states expressed concern when the principles of IHL were insufficient to protect civilians from the widespread use of anti-personal landmines and cluster munitions.¹⁰⁹ Like AWS, anti-personal landmines and cluster munitions are capable of lawful use, but in reality they were seldom used in conformity to the principles of IHL. States responded to the challenges raised by anti-personal landmines and cluster munitions through multilateral conventions rather than relying solely on the principles of IHL.¹¹⁰ It remains to be seen whether AWS will be as harmful to civilian populations as anti-personal landmines and cluster munitions have proven to be. Instead of waiting until the full ramifications of AWS are known, a proactive legal framework for their use should be developed.

Operational Solutions

Two operational solutions merit further discussion: (i) keeping humans in the loop to comply with IHL principles, and (ii) programming AWS in such a way as to avoid any potential violations of IHL.

Keeping humans in the loop

If humans were in the targeting decision loop, AWS would presumably be capable of adhering to the principles of IHL. For example, DoD Directive 3000.09 stipulates that “autonomous and semi-autonomous weapon systems shall be designed to allow commanders and operators to exercise appropriate levels of human judgment over the use of the force.”¹¹¹ However, this policy does not ensure that AWS will comply with IHL principles in the future.

Currently, the benefits of removing humans from the loop are minimal: the technology is not sophisticated enough to replace a human decision maker in most circumstances. As technology improves, the benefits of removing humans will grow. Those benefits include less expense, lower reliance on communication networks, and

¹⁰⁸ Schmitt, *supra* note 43 at 2.

¹⁰⁹ Hagger & McCormack, *supra* note 47 at 90.

¹¹⁰ See *Convention on the Prohibition of the Use, Stockpiling, Production and Transfer of Anti-Personal Landmines and on Their Destruction*, 18 September 1997, 2056 UNTS 211 (entered into force 1 March 2009); Diplomatic Conference for the Adoption of a Convention on Cluster Munitions, *Convention on Cluster Munitions*, 3 December 2008, CCM/77 (entered into force 1 August 2010).

¹¹¹ *DoD 3000.09*, *supra* note 3, s 4(a), at 2.

faster response times. Ensuring that humans stay in the loop would give AWS the ability to adhere to IHL principles, but it is not realistic as a permanent solution because states will be under pressure to lower operational costs and gain the competitive advantage by taking humans out of the equation.

Arkin's Ethical Governor

Arkin suggests it will be possible to embed ethics in AWS so as to ensure their compliance with IHL principles. By constraining the range of possible actions available to AWS through an "ethical governor" and "ethical behavioural control," Arkin contends that AWS will not only be capable of adhering to IHL, but that they will also be able to exceed human capabilities.¹¹²

Arkin's attempts to translate IHL principles into a logical and programmable structure offer practical solutions to some issues raised by AWS, but there are problems with this approach. Arkin's ethical governor assumes that "effective situational assessment methods exist" to ensure that AWS would not commit a lethal mistake.¹¹³ As demonstrated above, this assumption may not be appropriate. In addition, an ethical governor may only be implementable in larger systems because of the requirement to read and assess a complex environment.¹¹⁴

Keeping humans in the loop and implementing an ethical governor are practical solutions to the immediate challenges raised by the use of AWS. However, they do not address the fundamental concern raised by AWS: whether IHL principles can ensure that AWS are used legally.

Proposed legal solutions

Outright Ban

Human Rights Watch has proposed a blanket ban on autonomous robots, similar to anti-personal landmines. The plan would simply prohibit any use of lethal force by AWS without human input or supervision.¹¹⁵ If adopted, the ban would effectively ensure that AWS adhere to the principles of IHL.

In contrast, if the potential advantages of AWS are realized, an outright ban may actually undermine the principles of IHL. Assuming that AWS are better able to distinguish and act proportionally than humans, banning these weapons would increase the risks to civilians.¹¹⁶

Proponents of AWS argue that the mere possibility that they could be used illegally "is not a valid basis for imposing an across-the-board pre-emptive ban on the systems."¹¹⁷ An outright ban might realistically be a step too far, but its proposal does

¹¹² Ronald Arkin, "Governing Lethal Behavior: Embedding Ethics in a Hybrid Deliberative/Reactive Robot Architecture, Georgia Institute of Technology GVI Technical Report GIT-GVI-07-11" (2007) at 61, online: Georgia Institute of Technology: <<http://www.cc.gatech.edu>>. A full analysis of Arkin's proposal is outside the scope of this paper. For a more thorough description of Arkin's proposal, see Krishnan, *supra* note 5.

¹¹³ *Ibid* at 22.

¹¹⁴ Krishnan, *supra* note 5 at 109.

¹¹⁵ *Ibid* at 47.

¹¹⁶ Schmitt, *supra* note 43 at 25.

¹¹⁷ *Ibid* at 10.

enhance the discourse and promotes a serious consideration of the issues faced by implementing AWS.

Multilateral Conventions

There are many examples of new weapons technologies giving rise to multilateral conventions that restrain their use or development: biological weapons,¹¹⁸ chemical weapons,¹¹⁹ blinding lasers,¹²⁰ and landmines.¹²¹ These conventions employ a wide range of approaches to regulating the use of weapons. For example, the Biological Weapons Convention and Chemical Weapons Convention prohibits all state-parties from acquiring, developing, producing, stockpiling, or retaining biological or chemical weapons.¹²² In contrast, Protocol IV on Blinding Laser Weapons prohibits the use of permanently blinding laser weapons, but not their research or development.¹²³

Multilateral conventions have been successful in regulating these weapons because their use is widely recognized as contrary to the principles of IHL. The same cannot be said for AWS: there is still reasonable debate as to whether AWS will be capable of adhering to the principles of IHL. It has been suggested that a multilateral treaty that bans or limits the use of AWS is impractical. Consensus would be impossible to reach and compliance would not be guaranteed.¹²⁴

The annual meeting of state parties to the Convention on Certain Conventional Weapons (CCW) in November 2013 raised the topic of AWS. The state parties decided to establish an “informal Meeting of Experts” to “discuss the questions related to emerging technologies in the area of lethal autonomous weapons systems....”¹²⁵ The establishment of the four-day informal meeting was supported by the majority of state parties, including the US.¹²⁶ The proposal is a welcome recognition of the need to develop international standards and state commitments regarding the use of AWS. It remains to be seen whether CCW is the appropriate international convention to regulate the development and use of AWS because the weapons currently regulated under CCW, including blinding lasers and landmines, do not raise the same fundamental concerns.

Further development of AWS is necessary before a multilateral convention would be a viable option to regulate their use. However, waiting until the technology exists

¹¹⁸ *Convention on the Prohibition of the Development, Production and Stockpiling of Bacteriological (Biological) and Toxin Weapons and on their Destruction*, 10 April 1972, 1015 UNTS 163 (entered into force 26 March 1975).

¹¹⁹ *Convention on the Prohibition of the Development, Production and Stockpiling of Chemical Weapons and on their Destruction*, 13 January 1993, 1974 UNTS 45 (entered into force 29 April 1997).

¹²⁰ *Protocol to the Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons which may be deemed to be Excessively Injurious or to have Indiscriminate Effects (Protocol IV, entitled Protocol on Blinding Laser Weapons)*, 13 October 1995, 33 ILM 1218 (entered into force 30 July 1998).

¹²¹ *Protocol on Prohibitions or Restrictions on the Use of Mines, Booby-Traps and Other Devices (Protocol II) (As Amended on 3 May 1996)*, 10 October 1980, 2048 UNTS 133 (entered into force 3 December 1999).

¹²² Marchant, *supra* note 19 at 301.

¹²³ *Ibid* at 305.

¹²⁴ Anderson & Waxman, *supra* note 15 at 45.

¹²⁵ *Draft Final Report*, 2013 Session, Meeting of the High Contracting Parties to the Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to Be Excessively Injurious or to Have Indiscriminate Effects, (15 November 2013), CCW/MSP/2013/CRP at para 32, online: Reaching Critical Will <<http://www.reachingcriticalwill.org/images/documents/Disarmament-fora/ccw/MSP-2013/Documents/draftreport.pdf>>.

¹²⁶ *US Opening Statement at the Meeting of the Parties to the CCW*, (14 November 2013), online: Mission of the United States <<http://www.geneva.usmission.gov>>.

before implementing standards is not desirable. Guidelines and dialogue are essential during the development stage to ensure that AWS are engineered in a way that complies with IHL standards.

The Current Approach: The Obligation to Review and Codes of Conduct

Article 36 of Additional Protocol I places each member state “under an obligation to determine whether its employment would, in some or all circumstances, be prohibited by this Protocol or by any other rule of international law...”¹²⁷ Although the US is not a signatory to this Protocol, it has been suggested that they are obligated to review weapons and ensure that they are capable of complying with the principles of IHL.¹²⁸

It is current US policy to review weapons systems. The Department of Defense has prescribed that “the acquisition and procurement of DoD weapons and weapon systems shall be consistent with all applicable domestic law and treaties and international agreements..., customary international law, and the law of armed conflict...”¹²⁹ It has been proposed that this is enough to ensure AWS are used legally.¹³⁰

The states engaged in the development of AWS, particularly the US, have an inherent interest in regulating AWS. Anderson and Waxman argue the US must develop a set of principles to regulate and govern AWS for the benefit of the international community as a whole.¹³¹ There are multiple benefits to a code of conduct approach. Codes are more flexible than formal multilateral conventions, and are potentially more capable of adapting quickly to technological advances. They are relatively easy to create and adopt in comparison to multilateral conventions. However, there are disadvantages as well. Codes of conduct are not internationally binding and do not involve any international oversight.¹³² They do not offer long-term solutions, as more and more states become capable of developing and using AWS according to their own standards.

In addition, the obligation to review can be circumvented. The US Predator drone was reviewed when it was only capable of surveillance and had not yet been weaponized. Later, when it was equipped with Hellfire missiles, the US stated that they did not need to review the drone again, as both the drone and the missile had been reviewed separately.¹³³ A similar approach to AWS would not constitute a full review of AWS and should be avoided.

DoD Directive 3000.09 represents the current US Code of Conduct. It was released publicly, making it open to scrutiny and available for the benefit of the international community. Although it demonstrates American concern with the development of AWS, it does not go far enough in planning the future development of AWS. It relies on keeping humans in the loop, which may ensure the legality of AWS in the short term, but the Directive fails to prepare for the possibility that humans will be removed when it is beneficial and technologically possible to do so.

¹²⁷ *Additional Protocol I*, *supra* note 50, art 36.

¹²⁸ Schmitt, *supra* note 43 page 28.

¹²⁹ United States Department of Defence, *Directive Number 5000.01: The Defense Acquisition System*, 20 November 2007, online: Defense Technical Information Center <<http://www.dtic.mil/whs/directives/corres/pdf/500001p.pdf>>.

¹³⁰ Anderson & Waxman, *supra* note 15 at 48.

¹³¹ *Ibid* at 46.

¹³² Marchant, *supra* note 19 at 306.

¹³³ Human Rights Watch, *supra* note 2 at 23.

DoD policy may change over time to reflect the development of AWS. When it becomes advantageous and technologically possible to take humans out of the loop, a new policy will be implemented to ensure compliance with the principles of IHL. This approach may result in internal safeguards, but it does not necessarily promote international discussion or ensure compliance with IHL unless the new policy is also publicized. One country may adopt strict policies, while others may not. Relying on state-created codes of conduct and their obligation to review will result in unilateral decision-making. This would stymie the development of internationally recognized principles necessary to ensure the safe operation of AWS globally.

Framework Conventions

A final possibility for the regulation of AWS would be adoption of a framework convention.¹³⁴ This approach would combine the positive qualities of a multilateral convention and a more flexible ‘code of conduct’ approach. A minimalistic multilateral convention could be created that develops a process and institutional capacity to gradually develop a substantive legal regime. Examples of this type of convention include: the Vienna Convention for the Protection of the Ozone Layer,¹³⁵ the Framework Convention on Climate Change¹³⁶ and the Framework Convention on Tobacco Control.¹³⁷

A framework convention has several benefits. It acknowledges that a problem exists and draws the attention of experts and the public to the problem. It commits states to taking more substantive action in the future.¹³⁸ Although it would not be a clearly defined regime from the start, it would allow for the identification of the precise issues that need addressing. It would not commit states to a binding agreement before the full capabilities of AWS are known. Instead, it would allow for open dialogue amongst states that are at various stages of developing AWS.

A new framework convention is preferable to regulating the development of AWS under an existing convention, such as CCW. CCW currently regulates conventional weapons that have been used in combat and have had their capabilities of adherence to the principles of IHL examined over time. In contrast, AWS are undergoing rapid technological development and have not yet been deployed in military operations. The fundamental difficulty that AWS have in adhering to the principles of IHL must be addressed before they are operationalized, not after. As technology develops, the legal framework for AWS must be consistently and frequently re-examined. A framework convention can address the unique challenge raised by AWS: removing humans from warfare.

¹³⁴ *Ibid* at 313.

¹³⁵ *Vienna Convention on the Protection of the Ozone Layer*, 22 March 1985, 1513 UNTS 323 (entered into force 22 September 1988).

¹³⁶ *United Nations Framework Convention on Climate Change*, A/RES/48/189, UNGAOR, 48 Sess, Supp No 49, online: United Nations <<http://www.unhcr.org>> 2013.

¹³⁷ *World Health Organization Framework Convention on Tobacco Control*, 21 May 2003 (entered into force 27 February 2005), online: World Health Organization <<http://www.who.int>>.

¹³⁸ Marchant, *supra* note 19 at 314.

The Way Forward

A legal framework to ensure the lawful use of AWS must be developed as soon as possible. Some proponents of AWS suggest that it is too early to know how the technology will develop and that lawmakers must wait until a fully autonomous system is in hand before the legal questions can be resolved.¹³⁹ This view is not sustainable. AWS are not a technology of the distant future: their development is already underway. A legal framework should be created before the development of AWS advances to the point where their underlying architecture is difficult to change.

In *International Governance of Autonomous Military Robots*, the authors suggest that there are four requirements for a successful legal regime.¹⁴⁰ First, it must involve “clearly defined and articulated expectations” which identify precisely the problems that need to be addressed. Second, the solutions to those problems must be realistic and capable of actual implementation. Third, they must be “holistic and inclusive” and include all relevant stakeholders in discussions. Fourth, they must be “subject to assessment” that allows for improvement over time.¹⁴¹ While codes of conduct offer some advantages, a framework convention is best suited to meeting these four criteria. A framework convention can identify the precise problem of adherence to IHL that comes from removing humans from the loop. Its incremental approach allows for the realistic implementation of international standards and periodic reassessment. A framework convention, in contrast to a state-by-state soft law approach, allows for the inclusion of all stakeholders.

CONCLUSION

Although AWS are not yet a reality, it is likely that they soon will be. The perceived benefits of AWS ensure that their development will continue. This development will occur incrementally towards full autonomy. At first, humans will remain in the loop. As the technology progresses, human interaction with AWS will likely decrease until humans are out of the loop altogether. The removal of humans from the decision-making loop leads to profound concerns that AWS will be incapable of adhering to the principles of IHL.

This paper focused on two principles of IHL: distinction and proportionality. There are also concerns that AWS violate principles of command responsibility. This is an area that requires further consideration.¹⁴² The principles of distinction and proportionality both seem to require subjective human reasoning, creating challenges for AWS compliance.

The use of AWS in compliance with the principles of IHL will prove challenging. AWS will have to overcome issues of weak machine perception, the computational

¹³⁹ Anderson & Waxman, *supra* note 15 at 40.

¹⁴⁰ Marchant, *supra* note 19 at 272, n 1: The authors are “members of the Autonomous Robotics Thrust group of the Consortium on Emerging Technologies, Military Operations, and National Security...a multi-institutional organization dedicated to providing the basics for the ethical, rational, and reasonable understanding and management of the complex set of issues raised by emerging technologies...”

¹⁴¹ *Ibid* at 290-91.

¹⁴² See Sparrow, *supra* note 6. *Contra* Schmitt, *supra* note 43.

challenges of an open environment, and the potential of weak software. Their judgment must also meet the standard of a reasonable commander. AWS must be programmed in such a way that they are capable of quickly performing complex situational calculations to at least the level of a human soldier. Additionally, there are concerns that the technology will never improve as well as proponents suggest. There is also a concern that the use of AWS will lead to moral disengagement and the expansion of the battle space.

The operational solutions proposed for AWS are not enough to ensure meaningful compliance with IHL; instead, a legal regime is required. An outright ban would prevent any violations of IHL, but it is also unrealistic. A multilateral convention would bring stakeholders to the table and such conventions have been effective at regulating the use of weapons in the past. However, it may be too early in the development of AWS for states to reach a comprehensive agreement. Currently, individual states are creating codes of conduct for developing AWS. This is a practical short-term solution but does not adequately address the incremental development of AWS. A framework convention, which brings stakeholders to the table and enables the development of a multilateral convention over time, may offer the best solution to the regulation of AWS.

It is crucial that a framework is developed before AWS become common on the battlefield, and the momentum of weapons development stymies meaningful regulation. It is unlikely that Isaac Asimov's first law of robotics, that "[a] robot may not injure a human being..." has any place in the development of robotic weapons.¹⁴³ However, through inaction we risk dehumanizing warfare and suffering significant humanitarian consequences.

¹⁴³ Isaac Asimov, *I, Robot* (New York: Bantam Dell, 2004) at 37.