

January 18, 2013

Volume 17, Issue 4

The Law That Applies to Autonomous Weapon Systems

By Jeffrey S. Thurnher



ASIL Insights, international law behind the headlines, informing the press, policy makers, and the public.

Introduction

Even though most robotics experts predict fully autonomous weapon systems will not be available for use on the battlefield for a number of years,^[1] the debate over the lawfulness of such systems has nonetheless begun. In November 2012, Human Rights Watch released a report entitled “Losing Humanity: The Case against Killer Robots.” In the report, which was produced in association with Harvard Law School’s International Human Rights Clinic, Human Rights Watch asserts that fully autonomous weapons will be unable to comply with key law of armed conflict principles. Specifically, it claims that “such revolutionary weapons . . . would increase the risk of death or injury to civilians during armed conflict.”^[2] Accordingly, Human Rights Watch advocates for an international treaty prohibiting the development, production, and future use of fully autonomous systems.

Only a few days after the Human Rights Watch report was released, the U.S. Department of Defense (“DoD”) seemingly reached a different conclusion about the lawfulness of such weapons when it released DoD Directive 3000.09. Entitled “Autonomy in Weapon Systems,” the directive established policies and guidelines for the development of autonomous functions in weapons.^[3] Although the directive does not specifically call for the development of fully autonomous weapons, it provides the required approval chain and details the necessary legal reviews to be conducted before such systems may be designed. It further mandates the enactment of a series of precautionary measures, such as safety measures and anti-tampering mechanisms, designed to keep autonomous systems from striking unintended targets.

These two documents have brought to the forefront the underlying question of whether autonomous weapon systems will ever be able to fully comply with the tenets of the law of armed conflict. This *Insight* explores these broader legal concerns over autonomous weapons by articulating the relevant precepts of law. It is instructive, however, to first begin

RELATED ASIL INSIGHTS

[The International Law of Drones](#)

[The U.N. Arms Trade Treaty: Temporarily Holstered](#)

[Adoption of the Convention on Cluster Munitions](#)

[The Use of White Phosphorus Munitions by U.S. Military Forces in Iraq](#)

[Weapons of Mass Destruction and International Law](#)

[Outcome of the Sixth Review Conference of the Biological Weapons Convention, November-December 2006](#)

[The Treaty on the Non-Proliferation of Nuclear Weapons: Taking Stock after the May 2008 Preparatory Committee Meeting](#)

[Insights Archive](#)

DOCUMENTS OF NOTE

[Department of Defense Directive 3000.09, Autonomy in Weapon Systems](#)

[Rights Watch, Losing Humanity: The Case Against Killer Robots](#)

[Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of International Armed Conflicts](#)

[Tallinn Manual](#)

[HPCR Manual on International Law Applicable to Air and Missile Warfare](#)

[ASIL EISIL>>](#)

ORGANIZATIONS OF NOTE

[Human Rights Watch](#)

[Harvard Law School's International Human Rights Clinic](#)

[U.S. Department of Defense](#)

[International Committee of the Red Cross](#)

with a brief introduction of the specific technology involved with these systems.

The Context: Autonomous Technology and What It Means to Be Fully Autonomous

The emergence of autonomous technology represents a potential sea change for modern warfare. DoD defines autonomous systems as those that “once activated, can select and engage targets without further human intervention by a human operator.”^[4] Both Human Rights Watch and DoD acknowledge that such systems do not currently exist.^[5] However, in the past few years, developments in artificial intelligence have shown remarkable advances. Researchers are beginning to develop “novel algorithms,” which allow unmanned systems to operate without human intervention.^[6] Using so called “machine learning” and other innovative approaches, computer systems are now coming closer to replicating human thought processes.^[7] These systems are better able to decipher and determine answers to complex problems by taking approaches similar to the way humans learn by example.^[8]

As these technological advances have emerged, militaries have begun embedding many autonomous features into weapon systems.^[9] Human Rights Watch fears that this trajectory will ultimately lead to fully autonomous weapon systems designed to make lethal targeting decisions.^[10] It asserts that such autonomous weapon systems will be unable to comply with key law of armed conflict principles. Given that charge, it is essential to examine more closely the relevant law.

The Law: How to Determine the Lawfulness of a Weapon System

It is incontrovertible that the law of armed conflict applies to autonomous weapon systems. When determining the overall lawfulness of a weapon system, there are two distinct aspects of the law that need to be analyzed: weapons law and targeting law.^[11] The former verifies that the weapon itself is lawful. The latter determines whether the use of the weapon system during hostilities might be prohibited in some manner under the law of armed conflict. A weapon must satisfy both aspects before it may be lawfully used on a battlefield.

When analyzing whether the weapon system itself is lawful, there are two distinct rules that apply. The first rule is that the weapon system must not be indiscriminate by its very nature. A weapon is deemed indiscriminate by nature if it cannot be aimed at a specific target and would be as likely to strike civilians as combatants. Found in Article 51(4)(b) of Additional Protocol I to the Geneva Conventions,^[12] the rule is considered to be reflective of customary international law.^[13] Accordingly, all states, even those not a party to the Protocol (such as the United States), are bound to comply with this customary law rule against indiscriminate attack. The mere fact that an autonomous weapon system rather than a human might be making the final targeting decision would not render the weapon indiscriminate by nature. Instead, as long as it is possible to supply the autonomous system with sufficiently reliable and accurate data to ensure it can be aimed at a military objective, then the system would not be deemed indiscriminate by nature. In the end, any proposed autonomous weapon system must comply with this provision to be lawful.

The second rule, codified in Article 35(2) of Additional Protocol I, is that a weapon system cannot cause unnecessary suffering or superfluous injury.^[14] This rule, which is also reflective of customary international law,^[15] seeks to prevent needless or inhumane injuries to combatants. A classic example of an unlawful weapon under this rule is a warhead that is filled with glass. Such a warhead would unnecessarily complicate medical treatment and would consequently be unlawful. This rule only presents a problem for an autonomous system if the specific warheads or weapons installed on the system would violate the rule. The fact that the system autonomously decides to engage a target does not itself affect or violate the prohibition on unnecessary suffering or superfluous injury. To potentially be

The purpose of ASIL Insights is to provide concise and informed background for developments of interest to the international community. The American Society of International Law does not take positions on substantive issues, including the ones discussed in this Insight. Educational and news media copying is permitted with due acknowledgement.

The Insights Editorial Board includes: [Cymie Payne](#); [Tania Voon](#); and [David Kaye](#). Djurdja Lazic serves as the managing editor.

deemed lawful, a fully autonomous weapon system must only be armed with weapons and ammunition that comply with this rule.

To verify compliance with the two rules listed above, a state intent on fielding a new weapon must conduct a thorough legal review. This requirement for a legal review, which appears in Article 36 of Additional Protocol I,^[16] ensures that the weapon is not indiscriminate and that it would not cause unnecessary suffering or superfluous injury. The review also determines whether there is any other particular provision under the law of armed conflict which would prohibit the use of the weapon. Customary law requires this legal review of weapons and weapon systems, also referred to as the means of warfare, and these reviews are thereby required by all states, including those not a party to the Protocol.^[17] The DoD Directive sets forth the U.S. policy of requiring such a review at the early stages of development and again just prior to actually fielding the system.^[18] Furthermore, if a weapon system is significantly modified after its initial fielding, then an additional review would be necessary. The development of any fully autonomous weapon system would clearly require such legal reviews.

Assuming the particular weapon satisfies the above weapons law rules, the weapon must still be examined under targeting law to determine whether the actual use of the weapon might be prohibited in some manner. To conduct this analysis, three core law of armed conflict requirements are particularly salient: distinction, proportionality, and precautions in the attack. A weapon system, even one deemed lawful under the above tenets of weapons law, may not be lawfully used if, based on the circumstances, its use would violate any one of these three requirements.

The first requirement is distinction. Distinction is the most fundamental principle of the law of armed conflict. A customary law principle, distinction obliges a combatant to distinguish between combatants and civilians, as well as between military and civilian objects.^[19] The rule is also codified in Article 48 of Additional Protocol I, with companion rules in Articles 51 and 52.^[20] The principle is intended to protect the civilian population by directing military attacks against only military targets. The context and environment in which the weapon system operates play a significant role in this analysis. There may be situations in which an autonomous weapon system could satisfy this rule with a considerably low level ability to distinguish between civilian and military targets. Examples would include during high intensity conflicts against declared hostile forces or in battles that occur in remote regions, such as underwater, deserts, or areas like the Demilitarized Zone in Korea. At other times, such as on a complex counterinsurgency-type battlefield or in an urban environment, the demands on the systems to distinguish would be drastically higher. In the latter situations, even autonomous systems equipped with the most robust sensor packages may have difficulty fulfilling this requirement. Ultimately, for the use of an autonomous weapon system to be lawful, the system would be expected to reasonably distinguish between combatants and civilians (and between military objectives and civilian objects) given the particular environment and circumstances of the battlefield ruling at the time.

The second requirement, proportionality, requires combatants to examine whether the expected collateral damage from an attack would be excessive in relation to the anticipated military gain. This complex principle is reflective of customary international law^[21] and has traditionally involved a human judgment call evaluated on the basis of reasonableness. This rule is also codified in both Article 51(5)(b) and Article 57(2)(iii) of Additional Protocol I.^[22] Many critics of autonomous weapon systems, including Human Rights Watch, have questioned whether the systems can fulfill this requirement. To comply with the principle, autonomous weapon systems would, at a minimum, need to be able to estimate the expected amount of collateral harm that might come to civilians from an attack. Additionally, if civilian casualties were likely to occur, the autonomous systems would need to be able to

compare the amount of collateral harm against some predetermined military advantage value of the target. This step may present a significant challenge for autonomous weapon systems. The military advantage of a particular target is extremely contextual, and its value can change rapidly based upon developments on the battlefield. Human operators may be able to develop sliding scale-type mechanisms which regularly update and provide the autonomous weapon system with the relative military advantage value of a given target. Operators might also help fulfill this principle by detailing strict rules of engagement for these systems and establishing other controls, such as geographic or time limits on use. Nevertheless, these complicated issues would need to be resolved, if the future use of autonomous weapon systems is to comport with the principle of proportionality.

The third and final core requirement is the obligation to take feasible precautions in the attack. These precautions, which are customary in nature and codified in Article 57 of Additional Protocol I,^[23] present certain challenges to autonomous weapon systems. One such challenge will be the requirement to do everything feasible to verify that a target is a military one.^[24] Feasible, in this context, generally means “that which is practicable or practically possible, taking into account all circumstances prevailing at the time, including humanitarian and military considerations.”^[25] There may be times where the robust recognition capabilities of an autonomous system would be more precise (and, thus, more reliable) than a human in fulfilling this requirement. In other cases, depending on the circumstances (and what is practically possible), a force may have to augment the autonomous system with other sensors to help verify the target. Another significant challenge is the requirement to do everything feasible to choose a means of attack “with a view to avoiding, and in any event minimizing,” collateral damage.^[26] This precaution may, under certain circumstances, preclude the use of an autonomous system if a different type of a system would better protect civilians.^[27] With all of the required precautions in attack, there is inherently a value judgment about whether all feasible steps have been taken. How autonomous systems will reasonably make this value judgment may prove to be one of the biggest challenges in terms of compliance. Ultimately if a country intends to use an autonomous weapon system on a battlefield, it must ensure that the system can adequately take these feasible precautions.

Conclusion

In this debate, which is often infused with policy, morality, or ethical arguments, it is important for individuals in the international law field to distinguish those arguments from purely legal ones. The above rules outline the basic legal standards a weapon must meet to be deemed lawful. It remains to be seen whether in the future autonomous weapon systems will be banned or will flourish on the battlefield. Both the United States and groups like Human Rights Watch recognize the significance and the unique nature of these legal concerns, but, at least initially, they have staked vastly different approaches and positions towards handling autonomous weapon systems. The only thing clear at this point is that this legal debate will likely continue.

About the Author:

Jeffrey S. Thurnher is a Judge Advocate in the U.S. Army and a faculty member in the International Law Department at the U.S. Naval War College. Major Thurnher was the winner of the 2009 ASIL Lieber Society Military Prize. The opinions expressed herein are solely those of the author, and are not intended in any way to reflect an official position of the U.S. Naval War College, the U.S. Department of Defense, or any branch of the U.S. Government.

Endnotes:

- [1] Many experts have predicted that autonomous weapon systems will become the norm on the battlefield, but the expected timeline for that to happen is about twenty years. P.W. Singer, *Wired for War: The Robotics Revolution and Conflict in the 21st Century* 128 (2009).
- [2] Human Rights Watch, *Losing Humanity: The Case Against Killer Robots*, at 1 (Nov. 2012), available at http://www.hrw.org/sites/default/files/reports/arms1112ForUpload_0_0.pdf.
- [3] Department of Defense Directive 3000.09, *Autonomy in Weapon Systems* (Nov. 21, 2012), available at <http://www.dtic.mil/whs/directives/corres/pdf/300009p.pdf> [hereinafter DoD Directive 3000.09].
- [4] *Id.* at 13.
- [5] *Losing Humanity*, *supra* note 2, at 3; Department of Defense, DoD Directive 3000.09: *Autonomy in Weapon Systems: Response-to-Query Talking Points 1* (n.d.) (on file with author).
- [6] Colin Poitras, *Smart Robotic Drones Advance Science*, UConn Today (Oct. 4, 2012), <http://today.uconn.edu/blog/2012/10/smart-robotic-drones-advance-science/>.
- [7] Edward Larkin, *Siri, Watson, and Artificial Intelligence's Big Year*, London School of Economics Beaver Newspaper (Oct. 11, 2011), <http://edwardandlarkin.com/2011/10/11/siri-watson-and-artificial-intelligences-big-year/>.
- [8] Public Broadcasting Service [PBS], *Smartest Machines on Earth*, NOVA (Sept. 14, 2011), www.pbs.org/wgbh/nova/tech/smartest-machine-on-earth.html.
- [9] A Department of Defense ("DoD") scientific task force has even recommended that DoD "more aggressively use autonomy in military missions." Memorandum from Undersecretary of Defense for Acquisition, Technology, and Logistics, in Department of Defense, Defense Science Board, *The Role of Autonomy in DoD Systems* (July 2012), <http://www.acq.osd.mil/dsb/reports/AutonomyReport.pdf>. For a lengthier discussion of recent advancements in autonomous technology and uses by militaries, see Jeffrey S. Thurnher, *Legal Implications of Fully Autonomous Targeting*, 67 *Joint Force Q.* 77 (2012).
- [10] *Losing Humanity*, *supra* note 2, at 3.
- [11] Michael N. Schmitt, *Autonomous Weapon Systems and International Humanitarian Law: A Reply to the Critics*, Harv. Nat'l Sec. J. (forthcoming), available at <http://ssrn.com/abstract=2184826>. Targeting law is often also referred to as the rules that apply to the conduct of hostilities.
- [12] Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of International Armed Conflicts, art. 51(4)(b), June 8, 1977, 1125 U.N.T.S. 3 [hereinafter AP I].
- [13] The International Committee of the Red Cross ("ICRC") outlined which provisions of international law it deemed to be customary via an extensive study completed in 2005. It found Article 51(4)(b) of AP I to be reflective of customary law. International Committee Of The Red Cross, *Customary International Humanitarian Law rule 7* (Jean-Marie Henckaerts & Louise Doswald-Beck eds., 2005) [hereinafter Customary Law Study]. Although the United States did not accept the ICRC's determinations that certain provisions reflected customary international law (see Letter from John B. Bellinger, III, Legal Adviser, U.S. Dept. of State, and William J. Haynes, General Counsel, U.S. Dep't of Defense, to Dr. Jakob Kellenberger, President, ICRC (Nov. 3, 2006), available at http://lgdata.s3-website-us-east-1.amazonaws.com/docs/905/474175/US_LTR_on_customary_IHL_study.pdf), in practice, the United States has treated many of the provisions of AP I as legal obligations and has recognized them as reflective of customary international law. The provisions discussed throughout this *Insight*, unless otherwise stated, are ones that the United States appears to recognize as reflective of customary international law. For instance, with regard to Article 51(4)(b), the United States appears to consider the rule as reflective of customary international law given that it follows a similar rule under its treaty obligations from the Convention on Certain Conventional Weapons of 1980. See George Cadwalader, *The Rules Governing the Conduct of Hostilities in Additional Protocol I to the Geneva Conventions of 1949: A Review of Relevant United States References*, 14 *Y.B. Int'l Humanitarian L.* 133, 153 (2011).
- [14] AP I, *supra* note 12, art. 35(2).
- [15] Cadwalader, *supra* note 13, at 157; Customary Law Study, *supra* note 13, rule 70; Legality of the Threat or Use of Nuclear Weapons, Advisory Opinion, 1996 I.C.J. 226, ¶ 78 (July 8).

[16] AP I, *supra* note 12, art. 36.

[17] A legal review requirement is generally considered customary only with respect to the means of warfare. AP I Article 48 also requires a legal review of methods of warfare. An obligation to review new methods of warfare has not crystallized into customary international law. Tallinn Manual on the International Law Applicable to Cyber Warfare (Michael N. Schmitt gen. ed., forthcoming 2012) (commentary accompanying Rule 48).

[18] DoD Directive 3000.09, *supra* note 3, at 7-8.

[19] Cadwalader, *supra* note 13, at 157; Customary Law Study, *supra* note 13, rule 1; *Legality of the Threat or Use of Nuclear Weapons*, *supra* note 15, ¶¶ 78-79.

[20] AP I, *supra* note 12, arts. 49, 51-52.

[21] Cadwalader, *supra* note 13, at 157-58; Customary Law Study, *supra* note 13, rule 14.

[22] The rule specifies that an attack is indiscriminate if it is “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.” AP I, *supra* note 12, art. 51(5)(b).

[23] Cadwalader, *supra* note 13, at 161-62; Customary Law Study, *supra* note 13, rule 15.

[24] AP I, *supra* note 12, art. 57(2)(a)(i).

[25] Harvard Program on Humanitarian Policy and Conflict Research, *Manual on International Law Applicable to Air and Missile Warfare, With Commentary* 38 (2009).

[26] AP I, *supra* note 12, art. 57(2)(a)(ii).

[27] Conversely, there may be times when, under the circumstances, the use of an autonomous weapon system would be required, such as when its use is feasible and would offer greater protection to civilians.