INTERNATIONAL LAW STUDIES
PUBLISHED SINCE 1895
U.S. NAVAL WAR COLLEGE

Adapting the Law of Armed Conflict to Autonomous Weapon Systems

Kenneth Anderson, Daniel Reisner and Matthew Waxman

90 INT'L L. STUD. 386 (2014)

Volume 90 2014

Published by the Stockton Center for the Study of International Law

Adapting the Law of Armed Conflict to Autonomous Weapon Systems

Kenneth Anderson,[*] Daniel Reisner[**] and Matthew Waxman[***]

CONTENTS

I. INTRODUCTION

Weapon systems are becoming increasingly automated and arguably some autonomous military systems have been deployed for years. Recent advances in automated systems and the possibilities they portend have generated interest and anxiety within some militaries and defense ministries, and a movement of non-governmental activists seeking to ban fully autonomous weapons. In May 2014, the High Contracting Parties of the UN

* Professor of Law, Washington College of Law, American University; Member of the Hoover Institution Task Force on National Security & Law; Non-resident Senior Fellow, The Brookings Institution.

** Former head of the Israel Defense Force's International Law Department; currently the international law, defense, homeland security and aerospace partner at the law firm of Herzog, Fox & Neeman.

*** Liviu Librescu Professor of Law, Columbia Law School; Adjunct Senior Fellow, Council on Foreign Relations; Member of the Hoover Institution Task Force on National Security & Law.

Convention on Certain Conventional Weapons (CCW) convened an extensive discussion of the legal and ethical issues that autonomous weapons raise, while recognizing that many of these problems lie at an uncertain point in the future.

It is important that normative development regarding autonomous weapon systems head down a path that is coherent and practical.[1] By "autonomous weapon systems," we mean systems "that, once activated, can select and engage targets without further intervention by a human operator." We draw this definition from a 2012 U.S. Department of Defense policy directive, which remains the most extensive public pronouncement by any State on how it intends to proceed with regard to research, development and deployment of autonomous weapon systems.[2]

This paper addresses several questions that are critical to charting such a path. First, are autonomous weapon systems different from other new weapon systems, and, if so, how? Second, to the extent they are different, can and should autonomous weapon systems be regulated within the framework of the existing law of armed conflict? If yes, how should States go about doing so? If not, what alternative regulatory approach is appropriate?

We conclude that autonomous weapon systems have special features that pose risks and that create challenges in applying the existing law of armed conflict. Nevertheless, we conclude it is possible to adapt the existing framework to account for the features of autonomous weapons, and that the suggested alternative of prohibiting these systems outright is misguided. Instead, we propose a three-tiered process for regulating the development, deployment and use of autonomous systems.

---

1. This paper expands on arguments laid out in Kenneth Anderson & Matthew Waxman, *Law and Ethics for Autonomous Weapon Systems: Why a Ban Won't Work and How the Laws of War Can*, HOOVER INSTITUTION (Apr. 9, 2013), http://www.hoover.org/research/law-and-ethics-autonomous-weapon-systems-why-ban-wont-work-and-how-laws-war-can; also available as American University, WCL Research Paper 2013-11, Columbia Public Law Research Paper 13-351 (2013), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2250126.

2. U.S. Department of Defense, DoD Directive 3000.09, Autonomy in Weapon Systems 13 (2012) [hereinafter DoD Directive 3000.09].

II. THE INCREMENTAL DEVELOPMENT OF AUTONOMOUS WEAPON SYSTEMS

Autonomous weapon systems have existed in some sense for a long time. Many anti-personnel landmines and undersea mines could be thought of as autonomous in that once deployed they were configured (originally mechanically and later electronically) to detonate based on physical contact or other signals associated with targets. Many did so in very unsophisticated ways, unable to distinguish much more than pressure or movement, let alone select between lawful and unlawful targets. Conceptually, at least, they might fit the definition of autonomy if "select" is construed to mean merely "triggered" rather than "selection among" targets. This paper focuses instead, however, on technologically sophisticated systems in which capabilities for "selection among" is a specific design aim for the weapon. Some modern and very sophisticated autonomous (or at least very highly automated) weapon systems already exist. These are generally limited to use in defensive contexts against other machines and are deployed in environments such as the air or sea in which civilian risk is very small, and with respect to which human operators activate and monitor the system and can override its operation.

New autonomous weapon systems are gradually becoming incorporated into warfare as technology advances and capabilities increase one small automated step at a time. Increasing automation in weapons technology results from advances in sensor and analytical capabilities and their integration into—and in response to the increasing tempo of—military operations. It also results from political pressures to protect not just one's own personnel, but also civilian persons and property.[3] Although automation will be a general feature across battlefield environments and weapon systems, genuine *autonomy* in weapons will probably remain rare for the foreseeable future and be driven by special factors such as speed and the tempo required of particular kinds of operations.[4]

---

3. Kenneth Anderson & Matthew Waxman, Op-Ed, *Killer Robots and the Laws of War*, WALL STREET JOURNAL, Nov. 4, 2013, at A19, *available at* http://online.wsj.com/news/articles/SB10001424052702304655104579163361884479576.

4. *See* PETER SINGER, WIRED FOR WAR: THE ROBOTICS REVOLUTION AND CONFLICT IN THE 21ST CENTURY (2009). For a dissenting view that removing humans from military targeting is unlikely for the foreseeable future, see Werner J. A. Dahm, Op-Ed, *Killer Drones Are Science Fiction*, WALL STREET JOURNAL (Feb. 15, 2012), http://online.wsj.com/news/articles/SB10001424052970204883304577221590015475180.

"Automation" describes a continuum, and there are various ways to define places along it.[5] Terms like "semi-autonomous," "man-in-the-loop" and "man-on-the-loop" are used to convey different levels and configurations of machine-human interaction and degrees of independent machine decision making.[6] Rather than engaging in definitional debates, our analytical starting point is that new autonomous systems will develop incrementally as more functions, not just of the weapon but also of the platform (e.g., the vehicle or aircraft), are automated.

For example, intermediate levels of automation of weapon systems could include a robot that is pre-programmed to look for certain enemy weapon signatures and to alert a human operator of the threat, who then decides whether or not to pull the trigger. At a next level of automation, the system might be set with the human being not required to give an affirmative command, but instead merely deciding whether to override and veto a machine-initiated attack. Perhaps next the system would be designed to target and fire autonomously, but to wait and call for higher-level authorization only when it assesses possible collateral damage above a certain level. In some cases, a human operator might control only a single or very few sets of sensor and weapon units. In others, he might control or oversee an integrated network of sensor and weapon units. In still other cases, the move to automate the weapon system or even give it autonomy might be driven by automation of all the other non-weapon systems of the platform with which the weapon has to be coordinated. Eventually, these systems may reach the point of full autonomy for which, once activated, the human role is vanishingly small. As explained below, the tipping point from a highly-automated system to an "autonomous" one is very thin and, in practice, unstable.

Weapon systems that would be able to assess civilian status or estimate harm as part of their own independent targeting decisions, as in one of the above examples, do not exist today and research toward such capabilities currently remains in the realm of theory. Still, several modern highly-automated—and some would call them autonomous—weapon systems already exist. These are generally for use in battlefield environments such as naval encounters at sea where risks to civilians are small, and are general-

5. *See* DEFENSE SCIENCE BOARD, U.S. DEPARTMENT OF DEFENSE, TASK FORCE REPORT: THE ROLE OF AUTONOMY IN DOD SYSTEMS 3–5, 23–24 (2012).

6. *See* William C. Marra & Sonia K. McNeil, *Understanding "The Loop": Regulating the Next Generation of War Machines,* 36 HARVARD JOURNAL OF LAW AND PUBLIC POLICY 1139 (2013).

ly limited to defensive contexts against other machines in which human operators activate and monitor the system and can override its operation.[7] Examples include the U.S. Patriot and Phalanx anti-missile systems and Israel's Iron Dome anti-missile system.[8] Many more could lie ahead for a variety of battlefield environments and military missions in a future that is becoming less and less distant.

From a "demand-side" perspective, i.e., the needs of the military operator, increasing automation in weapons technology is an unsurprising response to the increasing tempo of military operations in which, other things being equal, the faster system wins the engagement. Automation permits military systems of all kinds (not just weapons) to act more quickly than people can assess, calculate and respond, and sometimes to act more precisely and accurately in responding to a military threat. Some of the earliest computational weapon systems emerged in World War II, such as the development of anti-aircraft guns that had rudimentary mechanisms for estimating trajectory and responding to enemy aircraft, targeting the anticipated path of the enemy bomber and firing in what, for that period, was high speed.[9]

Speed of response is a demand of military necessity, certainly, but it is not simply a question of winning the engagement. Speed through automation or autonomy can also serve to make the use of force in battle more precise. By shortening the time between the positive identification of a target and its attack, there is less likelihood that the situation might have changed, the target have moved or civilians have come into proximity. In the Libya hostilities in 2011, for example, NATO-manned attack aircraft were reportedly too slow and had too little loiter time to permit accurate targeting of highly mobile vehicles on the ground in an urban battlefield

7. *See* INTERNATIONAL COMMITTEE OF THE RED CROSS, REPORT OF THE ICRC EXPERT MEETING ON "AUTONOMOUS WEAPON SYSTEMS: TECHNICAL, MILITARY, LEGAL AND HUMANITARIAN ASPECTS," (May 9, 2014), *available at* http://www.icrc.org/eng/assets/files/2014/expert-meeting-autonomous-weapons-icrc-report-2014-05-09.pdf [hereinafter REPORT OF ICRC EXPERT MEETING].

8. U.S. Navy Fact File, *MK-15 Phalanx Close-In Weapons System (CIWS)*, http://www.navy.mil/navydata/fact_display.asp?cid=2100&tid=487&ct=2 (Nov. 15, 2013); Federation of American Scientists, *MK 15 Phalanx Close-In Weapons System (CIWS)*, (Jan. 9, 2003), http://www.fas.org/man/dod-101/sys/ship/weaps/mk-15.htm; Michael N. Schmitt & Jeffrey S. Thurnher, *Out of the Loop: Autonomous Weapon Systems and the Law of Armed Conflict*, 4 HARVARD NATIONAL SECURITY JOURNAL 231 (2013).

9. Another example is the Norden Bombsight, a basic analog computer used by the U.S. Air Force in WWII to greatly increase bombing accuracy.

with many civilians. In response, an appeal was made to the United States to supply, first, surveillance drones, and then armed drones that could speed up the targeting process.[10] Some version of this problem will drive demand for automation. On top of all this, there are always powerful pressures domestically and internationally on armed forces for advanced weapon technology to protect not just one's own personnel, but also civilian persons and property.

From a "supply-side" perspective, increasing automation—and eventually autonomy—of weapon systems grows from ever-continuing advances in sensor and analytic technologies, machine-learning and their fusion. Importantly, development of many of the enabling technologies of autonomous weapon systems—artificial intelligence and robotics, for example—are being driven by private industry for many commercial and societally-beneficial purposes (consider self-driving cars, surgical robots, and so on). They are developing and proliferating rapidly, independent of military demand and investment.[11] Such civilian automated systems are already making daily decisions that have potential life and death consequences, such as aircraft landing systems. While most people are generally aware that these types of systems are highly automated (or even autonomous for some functions), and have become wholly comfortable with their use, relatively little public discourse has addressed the increasing decision-making role of autonomous systems in potentially life-threatening situations. To the extent that such automation and robotics technologies come to be widely understood as more effective, safe and reliable than human judgment in many non-military realms, their use will almost certainly migrate into military ones. Indeed, future generations that perhaps come to routinely trust the computerized judgments of self-driving vehicles are likely to demand, as a moral matter, that such technologies be used to reduce the harms of war. It is largely a question of whether such systems work or not, and how well.

---

10. *See*, *e.g.*, Julian E. Barnes, *U.S. Launches Drone Strikes in Libya,* WALL STREET JOURNAL, Apr. 22, 2011, at A6, *available at* http://online.wsj.com/news/articles/SB10001424052748704889404576277413211029304 ("Allied aircraft have proved unable to fully protect civilians and rebel forces under heavy attack from Gadhafi loyalists. The drones can get closer, allowing more-precise strikes, and loiter longer, giving them a better chance of finding hidden targets. . . . Drones have been used for reconnaissance missions from the start of the conflict, but in recent days, NATO commanders had asked the U.S. to provide armed Predator strikes.").

11. *See* ROBERT O. WORK & SHAWN BRIMLEY, 20YY: PREPARING FOR WAR IN THE ROBOTIC AGE 23–27 (2014).

No doubt, the use and especially proliferation of autonomous weapons will pose significant risks and challenges for law and regulation. We address below some legal objections to autonomous weapons as such. But it bears recognizing at the outset that, as with any technologically advanced weapon system, there are risks and dangers that include machine malfunction and machines whose design (hardware, software or interaction with human operators) underperforms a legally essential task or produces unpredictable effects (including when autonomous weapon systems engage each other or not-fully predictable machine decisions resulting from probabilistic programming or machine (self-) learning).[12] Beyond issues of the autonomous weapon system itself, strategic or political issues include concerns about a State armed with autonomous weapons being too willing to resort to military force because these weapons might reduce perceived risks to a side's own soldiers or to civilians.[13] Moreover, these systems might be thought to

12. For an introduction to robotics using machine self-learning and probabilistic programming, see Ryan Calo, *Robotics and the New Cyberlaw*, 103 CALIFORNIA LAW REVIEW 101, 125–26 (forthcoming 2015), advance draft *available at* www.robotics businessreview.com/pdfs/roboticscyberlaw.pdf (describing how machine self-learning might produce machine behaviors that cannot be fully predicted in advance). For a technical introduction to the probabilistic programming or machine learning utilized in, for example, Google's self-driving cars, see SEBASTIAN THRUN, WOLFRAM BURGARD & DIETER FOX, PROBABILISTIC ROBOTICS (2005). For arguments, *contra*, that autonomous weapon systems will never in fact be able to perform adequately because of limitations of the ability to program their behaviors, see, e.g., Noel E. Sharkey, *The Evitability of Autonomous Robot Warfare*, 94 INTERNATIONAL REVIEW OF THE RED CROSS 787 (2012), *available at* http://www.icrc.org/eng/resources/documents/article/review-2012/irrc-886-sharkey.htm. For arguments that as a matter of fundamental morality targeting and engagement decisions should not be entrusted to robots, and raising concerns about accountability for mistake, errors and design flaws, see, e.g., Peter Asaro, *On Banning Autonomous Weapon Systems: Human Rights, Automation, and the Dehumanization of Lethal Decision-Making*, 94 INTERNATIONAL REVIEW OF THE RED CROSS 687 (2012), *available at* http://www.icrc.org/eng/resources/documents/article/review-2012/irrc-886-asaro.htm; WENDELL WALLACH & COLIN ALLEN, MORAL MACHINES: TEACHING ROBOTS RIGHT FROM WRONG 47–48 (2010).

13. Autonomous weapon systems and remotely-piloted armed drones raise the same issue: whether the features of the weapon system that limit risk to soldiers and make it more discriminating with respect to civilians are, ironically, the very features that make it, on some views, not just easier for a party to resort to force, but "too easy" to do so. *See, e.g.*, UK MINISTRY OF DEFENCE, JDN 2/11, THE UK APPROACH TO UNMANNED AIRCRAFT SYSTEMS ch. 5 (2011); MEDEA BENJAMIN, DRONE WARFARE: KILLING BY REMOTE CONTROL 124–43 (rev. ed. 2013). For a response, see Kenneth Anderson, *Efficiency* in Bello *and* ad Bellum*: Making the Use of Force Too Easy?*, *in* TARGETED KILLINGS: LAW AND MORALITY IN AN ASYMMETRICAL WORLD 374 (Claire Finkelstein, Jens David Ohlin

undermine individual disciplinary and accountability systems in the law of armed conflict.[14]

Note that many of these of these concerns—machine malfunction, marginal diminution of some political or ethical constraints on using force, and misuse or abuse—are not unique to autonomous weapons. They are true of many military technologies and long-standing targeting practices, including artillery, stand-off manned aircraft, missiles, rockets and other over-the-horizon weaponry. None of these weapon systems is regarded today as inherently illegal; the question is one of lawful use.

Autonomous weapons also offer important potential benefits though, not only in terms of military effectiveness, but in terms of humanitarian protection as well. Existing systems mentioned earlier help protect friendly forces and populations from modes of attack that are too fast or complex for human reaction and decision making. Like remotely piloted vehicles, some autonomous weapon systems can operate without exposing personnel to the direct risk of enemy fire. This, in turn, reduces pressures on combatant personnel to resort to greater force to address threats, with the possibility of greater harms resulting to civilians and civilian objects.

Autonomous weapon systems may also reduce risks to civilians by making targeting more precise and firing decisions more controlled, especially compared to human-soldier failings that are so often exacerbated by panic, vengeance or other emotions, as well as the limits of human senses and cognition. Empathy and sympathy are among the better angels of human nature, but one of Human Rights Watch's significant claims in *Losing Humanity*—that a fundamental objection to autonomous weapon systems is that they take these emotions out of battlefield targeting and firing decisions—flies in the face of how much of the structure of the law of armed conflict exists to address, imperfectly, the effects of human soldiers' battlefield emotions, starting with fear, anger and vengeance, exacerbated under conditions of hunger, exposure, uncertainty and so on.[15] The International Committee of the Red Cross (ICRC) has taken a sensibly cautious approach to this question, observing in a 2011 report that "emotion, the loss

---

& Andrew Altman eds., 2012), *available* as *Efficiency* in Bello *and* ad Bellum*: Targeting Killing Through Drone Warfare at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1812124.

14. *See*, *e.g.*, HUMAN RIGHTS WATCH, LOSING HUMANITY: THE CASE AGAINST KILLER ROBOTS sec. VI (Problems of Accountability for Fully Autonomous Weapons) (2012), *available at* http://www.hrw.org/reports/2012/11/19/losing-humanity-0 [hereinafter LOSING HUMANITY].

15. *Id.* at 37–38.

of colleagues and personal self-interest is not an issue for a robot, and the record of respect for IHL by human soldiers is far from perfect, to say the least."[16] Weapon systems with greater and greater levels of automation could—at least in some battlefield contexts—reduce misidentification of military targets, better detect or calculate possible collateral damage, or allow for using a smaller quanta of force compared to human decision making.[17]

That is not to say that autonomous systems do not raise some difficult issues. Perhaps foremost among these is the fact that as machine-learning and artificial intelligence technologies develop, it is becoming increasingly clear that human beings may not necessarily always be able to understand how (and possibly why) autonomous systems make decisions. In some cases, this is a question of the complexity of the system, making it realistically impossible to fully predict in advance how the system will behave, to understand the reasoning behind the system's decision making as it happens or even to reconstruct after the fact how it did behave. In other cases, the nature of the programming is probabilistic, making it impossible to know with complete certainty what the machine's decision will be.[18] This realization may lie at the core of many of the fears and concerns that have fueled the campaign to ban autonomous systems, discussed further below. But are autonomous weapon systems truly "game changers," with respect to the seeming loss of direct human control over weapons and their use on the battlefield?

We do not believe they are. Even today, operators of highly-automated military equipment, similar to users of civilian technology, quite often have very little idea as to exactly how the technology they are using functions internally as an electronic system, nor even the logical process of how the system reaches results, conclusions or decisions. And yet, the fact that we do not understand the internal mechanics (or electronics or software pro-

16. INTERNATIONAL COMMITTEE OF THE RED CROSS, INTERNATIONAL HUMANITARIAN LAW AND THE CHALLENGES OF CONTEMPORARY ARMED CONFLICT: REPORT PREPARED FOR THE 31ST INTERNATIONAL CONFERENCE OF THE RED CROSS AND RED CRESCENT 40 (2011), *available at* http://www.icrc.org/eng/resources/documents/report/31-international-conference-ihl-challenges-report-2011-10-31.htm [hereinafter ICRC 31ST CONFERENCE REPORT].

17. *See* Shane R. Reeves & William J. Johnson, *Autonomous Weapons: Are You Sure These Are Killer Robots? Can We Talk About It?*, ARMY LAWYER, Apr. 2014, at 25–26.

18. *See* Calo, *supra* note 12, at 125–26 (discussion of "emergent behavior" and machine learning enabling the machine to self-improve, but with possibly unpredictable learning paths and results).

gramming) of a system does not mean that we are unable to use such a system within our existing rule structures. It also does not mean that the humans operating the equipment do not have a general understanding of the operating parameters of the system and its general capabilities and limitations. As will be discussed in further detail below, the law's focus has always been—and should continue to be—on (a) ensuring that a weapon system is not inherently unlawful and (b) ensuring that the individual soldier or commander deploying or activating the lawful weapon system does so in accordance with the laws of warfare, to include ensuring that such systems are not deployed in situations for which they were not designed and which could cause unnecessary or unjustified harm. This important relationship between the weapon system and its human operator will run true for autonomous systems as well.

III. MISGUIDED CALLS TO PROHIBIT AUTONOMOUS WEAPONS PER SE

Some critics argue that legal and ethical deficiencies of autonomous weapons—or the policy tendencies that autonomous weapon capabilities would unleash—demand that such systems be prohibited. The assumptions behind calls to ban autonomous weapon systems and the form of proposed bans vary.

Some doubt that technology can ever be good enough to make sufficiently precise decisions to meet the legal and ethical requirements of distinction or proportionality, which are discussed below.[19] Others believe that these legal or ethical principles inherently require human judgment—that lethal targeting is or should always be guided directly by a moral agent who can be held accountable for culpable failures.[20] And others acknowledge that though autonomous weapon systems may not be unlawful under existing law, especially since technological development is uncertain and what costs and benefits any particular system might offer is today merely speculative, they nevertheless should be banned (at least for the foreseeable future, pending thorough international discussion and agree-

19. *See* LOSING HUMANITY, *supra* note 14, at 30–35; Noel Sharkey, Op-Ed, *America's Mindless Killer Robots Must Be Stopped*, GUARDIAN (Dec. 3, 2012), http://www.theguardian.com/commentisfree/2012/dec/03/mindless-killer-robots.

20. *See*, LOSING HUMANITY, *supra* note 14, at 35–36; Asaro, *supra* note 12.

ment as to their proper legal framework) as a prophylactic precautionary measure because of risks they pose, both known and unknown.[21]

One way to ban—or to interpret existing law as banning—autonomous weapons is to define some maximum level of autonomy for any weapon system and prohibit any machine system that exceeds it. A variant approach is to define some minimum legal level of human control. Human Rights Watch, for example, has called for a preemptive "ban [on] fully autonomous weapons," which "should apply to robotic weapons that can make the choice to use lethal force without human input or supervision." It also proposes to ban their "development, production, and use," as well as calling for "reviews" of "technologies and components that could lead to fully autonomous weapons."[22] The International Committee for Robot Arms Control, an organization dedicated to reducing threats from military robotics, calls for the "prohibition of the development, deployment and use of armed autonomous unmanned systems."[23] A British nongovernmental organization dedicated to the regulation of certain weapons argues that lethal decision making should require "meaningful human control."[24] This idea of requiring a minimum level of "meaningful human control" emerged as a major theme in discussions among States and advocacy groups at the 2014 UN CCW meeting.[25] Instead of encouraging a perma-

---

21. Special Rapporteur on Extrajudicial, Summary or Arbitrary Executions, *Report*, 20–21*,* U.N. Human Rights Council, U.N. Doc. A/HRC/23/47 (Apr. 9, 2013) (by Christof Heyns) [hereinafter Report of the Special Rapporteur]. It appears that for some in the autonomous weapons debate, either a ban or moratorium is a virtuous application of some form of the precautionary principle—generically, a policy heuristic of "first, do no harm" (i.e., prefer little risk of harm or reduction of risk even if it means giving up large potential benefits). *See*, *e.g.*, Brian Rappert et al., *The Roles of Civil Society in the Development of Standards around New Weapons and Other Technologies of Warfare*, 94 INTERNATIONAL REVIEW OF THE RED CROSS 765, 767 (2012) (endorsing "putting in place so-called precautionary measures . . . such as initially limiting use").

22. LOSING HUMANITY, *supra* note 14, at 46–47.

23. *Original Mission Statement*, ICRAC, http://icrac.net/statements/ (last visited July 29, 2014).

24. Memorandum from Article 36 for Delegates to the Convention on Certain Conventional Weapons (CCW), Article 36: Structuring Debate on Autonomous Weapon Systems (Nov. 2013), http://www.article36.org/wp-content/uploads/2013/11/ Autonomous-weapons-memo-for-CCW.pdf.

25. *See* Chairperson of the Meeting of Experts, UN Convention on Certain Conventional Weapons, Report of the 2014 Informal Meeting of Experts on Lethal Autonomous Weapons Systems (LAWS) ¶ 20, http://www.unog.ch/80256EDD006B8954/%28%20httpAssets%29/350D9ABED1AFA515C1257CF30047A8C7/$file/Report_AdvancedVersion_10June.pdf (last visited July 29, 2014).

nent ban on autonomous weapons, Christof Heyns, the UN Special Rapporteur on Extrajudicial, Summary or Arbitrary Executions, has proposed a moratorium, calling for "all States to declare and implement national moratoria on the testing, production, assembly, transfer, acquisition, deployment and use of [lethal autonomous robotics] until such time as an internationally agreed upon framework . . . has been established."[26]

Before addressing some of the enforceability problems and dangers of any effort to prohibit autonomous weapons, it is important to understand the proposed formulas do not, as it may seem initially, contain a bright line that would be useful for promoting adherence. Lawyers experienced in the law of armed conflict will quickly see that each of these seemingly clear-cut definitions leaves many open questions as to what systems would be banned under any particular formulation. Even something as seemingly plain as "lethal decision making" by a machine does not address, among other things, the lawfulness of targeting a tank, ship or aircraft which is ultimately the source of the threat, but inside of which is a human combatant.

Beyond definitions, the technology and basic architecture of an autonomous system and a nearly autonomous, highly automated system are basically the same—if you can build a system that is nearly autonomous, for example with human override, then you can probably reprogram it to eliminate that human role. Moreover, whether a highly automated system—say, one with a human supervisor who can override proposed firing decisions—is in practice operating autonomously depends on how it is being manned, how operators are trained and how effectively oversight is exercised. It also depends on operational context and conditions, which may limit the degree to which the human role is in any way meaningful.

For these and other reasons, a fully autonomous system and a merely highly-automated system will be virtually indistinguishable to an observer without knowing a lot about how that system is used in particular operational conditions. The difference might not matter very much in practice, given the variable performance of human operators. In any case, these systems will be easily transitioned from one to the other. The blurriness of these lines means that it will be very difficult to draw and enforce prohibitions on "fully" autonomous systems or mandates for minimum levels of human decision making. Given the great practical difficulty of distinguishing between autonomous and highly automated systems, applying a legal

26. Report of the Special Rapporteur, *supra* note 21, at 21.

ban on autonomous systems would be relatively easy to circumvent and very difficult to enforce.

At the same time, and as alluded to above, imposing a general ban on autonomous systems could carry some highly unfavorable consequences—and possibly dangers. These could include providing a clear advantage in autonomous weapon technology to those States which generally would not join (or in reality comply with) such a ban. They could also include losing out on the numerous potential advantages of autonomous systems of improving decision making on the battlefield, including through avoiding emotion-based response; improving system accuracy, thereby probably minimizing collateral injuries; and possibly limiting human loss of life on both sides and among civilians.

IV. REGULATING THE *USE* OF AUTONOMOUS WEAPONS UNDER THE LAW OF ARMED CONFLICT

Calls for a prohibitory ban are usually based on creating new treaty law. Without a new treaty, however, any new weapon system must still comply in its design and usage with the law of armed conflict. Article 36 of Protocol I, for example, requires that "in the study, development, acquisition or adoption of a new weapon, means or method of warfare, a High Contracting Party is 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 applicable to the High Contracting Party."[27] Additionally, assuming a weapon is not per se illegal, its operational uses must also comply with the law of targeting.

27. Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of International Armed Conflicts art. 36, June 8, 1977, 1125 U.N.T.S. 3 [hereinafter API]. Article 36 is widely regarded among scholars as expressing customary law with respect to "legal reviews of new *means* of warfare before their use is generally considered," but such consensus is "lacking as to whether an analogous requirement exists to perform legal reviews of new *methods* of warfare." Means of warfare refers to "weapons and weapon systems," whereas methods refer to the "tactics, techniques, and procedures" by which hostilities are conducted. *See* Schmitt & Thurnher, *supra* note 8, at 271. Given that neither the United States nor Israel is a party to Protocol I, and given that the authors include former national security government officials of each, we believe it prudent to observe that neither State appears to have expressed an official legal view as to the customary law status of each aspect of Article 36; however, the military establishments of both States have long maintained highly-developed and formalized processes for reviewing the legality of weapons. The processes are similar to what Article 36 requires of State parties to Protocol I.

The legality of the weapon itself turns on three basic rules.[28] First, the weapon system cannot be indiscriminate by nature.[29] This is not a question of whether there might be circumstances—the manner of use and battlefield conditions—in which the weapon could not be aimed in a way to comply with the legal requirement of "distinction" between lawful military targets and civilians. That would be true of nearly any weapon, because any weapon could be deliberately misused. Rather, the rule runs to the nature of the weapon *in the uses for which it was designed*[30] or, as some authorities have put it, its "normal" uses, i.e., the uses for which it was intended.[31] As a mat-

28. *See*, *e.g.*, Judge Advocate General, U.S. Air Force, AFI51-402, Legal Reviews of Weapons and Cyber Capabilities 3.1.1, 3.1.2 (2011) (Technical legal guidance for the legal review of weapon systems, specifying the scope of legal review of weapons to encompass both specific treaty rules and customary law rules regarding the use of the weapon, and beyond that to consider, even in the absence of an express prohibition, whether the weapon is "of a nature" to inflict superfluous injury or unnecessary suffering upon combatants and whether it is "capable of being directed against a specific military objective and, if not, is "of a nature to cause an effect on military objectives and civilians or civilian objects without distinction.").

29. We leave aside here a number of technical complications. For example, notwithstanding the use of "of a nature" by the U.S. Air Force weapons review guidance with respect to indiscriminate weapons, the language "of a nature" appears in Article 35(2) of the Additional Protocol I to the Geneva Conventions, with reference not to indiscriminate weapons with respect to civilians, but only weapons of a nature to cause superfluous injury or unnecessary suffering to combatants. *See* API, *supra* note 27, art. 35(2). There is reasonably widespread State practice and agreement that a similar concept applies to indiscriminate weapons, with, however, considerable debate over the question of how this "nature" is to be determined—whether by reference to a weapon's design, "normal" purpose and use, etc. *See* WILLIAM H. BOOTHBY, WEAPONS AND THE LAW OF ARMED CONFLICT 69–85 (2009). In general, we have followed the approach taken by both Boothby, and Schmitt and Thurnher.

30. It bears noting that even the ban on blinding laser weapons—the leading example of a ban on a technologically advanced weapon that had not yet been deployed—is limited to "laser weapons specifically designed, as their sole combat function or as one of their combat functions, to cause permanent blindness to unenhanced vision." Protocol on Blinding Laser Weapons art. 1, Oct. 13, 1995, 1380 U.N.T.S. 370. It underscores the general rule that the lawfulness of a weapon as such turns on the use for which it was designed or intended.

31. YORAM DINSTEIN, THE CONDUCT OF HOSTILITIES UNDER THE LAW OF ARMED CONFLICT 62 (2d ed. 2010) ("It is necessary to differentiate between (a) weapons that are employed in specific circumstances contrary to the principle of distinction . . . and (b) weapons that by their very nature or design cannot possibly maintain the distinction in any set of circumstances. The fact that certain weapons are used indiscriminately in a particular military engagement does not stain the weapons themselves with an indelible imprint of illegality . . . .").

ter of grounds to find a weapon unlawful, this is a tough standard and very few weapons are illegal per se because they are indiscriminate by nature. Rather, legal weapons are used in an indiscriminate manner—a serious violation of the law of armed conflict, certainly, but running to the actual use of the weapon.[32]

Second, a lawful weapon system cannot be "of a nature" to cause "unnecessary suffering or superfluous injury."[33] This provision aims to protect combatants from needless or inhumane suffering, such as shells filled with glass shards that would not be detectable by an x-ray of the wound.[34] It is a rule that applies solely to combatants, not civilians (who are protected by other law of armed conflict provisions).[35] Like the "indiscriminate by nature" rule, it sets a high bar; unsurprising, given the forms of violence that can lawfully be inflicted upon combatants in armed conflict.

Third, a weapon system can be deemed illegal per se if the harmful effects of the weapon are not capable of being "controlled."[36] The rule against weapons with uncontrollable harmful effects is paradigmatically biological weapons, in which a virus or other biological agent cannot be controlled or contained; once released, it goes where it goes. Once again, even though many rules of the law of armed conflict prevent the use of weapons in circumstances that might have uncontrolled effects, the bar to make the weapon itself illegal per se is high.

None of these rules renders a weapon system illegal per se solely on account of it being autonomous. If a fully autonomous weapon system were supplied with sufficiently reliable parameters and it were able to act on them so as to be able to strike specific targets on the same legal terms of

32. Schmitt & Thurnher, *supra* note 8, at 276 (weapons reviews "examine only the legality of a weapon system as such, not its use in any particular circumstance.").

33. API, *supra* note 27, art. 35(2).

34. For more examples, see COMMENTARY ON THE ADDITIONAL PROTOCOLS OF 8 JUNE 1977 TO THE GENEVA CONVENTIONS OF 12 AUGUST 1949, ¶ 1419 (Yves Sandoz, Christophe Swinarski & Bruno Zimmermann eds., 1987).

35. Schmitt & Thurnher, *supra* note 8, at 244.

36. API, *supra* note 27, art. 51(4)(c); Schmitt & Thurnher, *supra* note 8, at 250 ("[Article 51(4)(c)] reflects customary international law. It disallows weapon systems that, despite being able to strike their targets accurately, have uncontrollable effects."). Because the rule concerns "effects," the claim that some might make is that autonomous weapon systems are "uncontrollable" because for a weapon system equipped with sophisticated probabilistic programming not every decision taken by the machine would be predictable in advance, thus would be by definition uncontrolled. But apart from other aspects of control of the machine, uses and operations, the rule is about effects that cannot be uncontrolled, not an uncontrolled weapon.

discrimination that would apply to a human soldier, that the weapon system was "autonomous" would not violate the "indiscriminate by nature" rule. Although some might view an autonomous weapon system as "uncontrollable," its *effects* are not uncontrollable within the meaning of the legal provision.[37]

Even if a weapon system is not per se illegal, it might still be prohibited in some—even many—battlefield environments, or in particular uses on a particular battlefield. But in other circumstances, of course, the weapon might also be legal. As the ICRC put the point in 2011 with respect to new weapon technologies generally, the question is not whether the "new technologies are good or bad in themselves, but instead what are the circumstances for their use."[38]

Targeting law governs the circumstances of the use of lawful weapons and includes three fundamental rules: discrimination (or distinction), proportionality and precautions in attack.[39] Distinction requires that a combatant, using reasonable judgment in the circumstances, distinguish between combatants and civilians, as well as between military and civilian objects.[40] Although use of autonomous weapon systems is not illegal per se, a requirement for their lawful use—the ability to distinguish lawful from unlawful targets—might vary enormously from one weapon system's technology to another. Some algorithms, sensors or analytic capabilities might perform well, others poorly.

Moreover, these capabilities are measured against particular battlefield environments; the "context and environment in which the weapon system operates play a significant role in this analysis."[41] Air-to-air combat between military aircraft over the open ocean, for example, might one day take place between autonomous systems, as the technological pressures for greater

37. Schmitt and Thurnher are right to conclude that the "likelihood of an autonomous weapon system being unlawful per se is very low" on any of the three grounds described above. Schmitt & Thurnher, *supra* note 8, at 250.

38. ICRC 31ST CONFERENCE REPORT, *supra* note 16, at 40.

39. See IAN HENDERSON, THE CONTEMPORARY LAW OF TARGETING: MILITARY OBJECTIVES, PROPORTIONALITY, AND PRECAUTIONS UNDER ADDITIONAL PROTOCOL I 23–41 (2009).

40. API, *supra* note 27, art. 48 ("Parties to the conflict shall at all times distinguish between the civilian population and combatants and between civilian objects and military objectives . . . .").

41. Jeffrey S. Thurnher, *The Law that Applies to Autonomous Weapon Systems*, 17 AMERICAN SOCIETY OF INTERNATIONAL LAW INSIGHTS (Jan. 18, 2013), www.asil.org/insights/volume/17/issue/4/law-applies-autonomous-weapon-systems.

speed, ability to endure torque and inertial pressures, and so on, finally result in an unmanned craft. As the speed of flight and every other aircraft system becomes faster to the point that only an autonomous system can fly the plane, at that point and at those speeds the weapon systems may also have to be automated in order to operate effectively with the rest of the aircraft systems.

Distinction is highly unlikely to be an issue in that particular operational environment, however, because the combat environment would be lacking in civilians. As Jeffrey Thurnher points out, 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."[42] Yet there would be many operational environments in which meeting the requirements of distinction by a fully autonomous system would be very difficult—urban battlefield environments in which civilians and combatants are commingled, for example. This is not to say that autonomous systems are thereby illegal. Quite the opposite, in some settings their use would be legal and in others illegal, depending on how technologies advance.

Proportionality requires that the reasonably anticipated military advantage of an operation be weighed against the reasonably anticipated civilian harms.[43] As with the principle of distinction, there are operational settings—air-to-air combat over open water, tank warfare in remote uninhabited deserts, ship antimissile defense, undersea anti-submarine operations, for example—in which civilians are not likely to be present and which, in practical terms, do not require weighing military advantage against civilian harms. Conversely, in settings such as urban warfare, proportionality is likely to pose very difficult conditions for machine programming, and it is widely recognized that whether and how such systems might one day be developed is simply an open question.[44] To be fair, many military lawyers have questioned whether human soldiers are capable of truly applying this ambiguous test either.

Programming proportionality is not simply a question of sensors or computational sophistication in order to reasonably identify a person as a lawful target in the sense of distinction. While everyone agrees that civilian harm should not be excessive in relation to military advantages gained, the

42. *Id.*

43. The customary law rule is codified in Additional Protocol I. API, *supra* note 27, arts. 51(5)(b), 57(2)(a)(iii).

44. *See* ICRC 31ST CONFERENCE REPORT, *supra* note 16.

comparison is apples and oranges. Although there is a general sense that such excess can be determined in truly gross cases, there is no accepted formula that gives determinate outcomes in specific cases.[45] Some military lawyers proceed largely casuistically, building on what was done in prior situations and examining similarities and differences.[46] Difficult or not, proportionality is a fundamental requirement of the law and any completely autonomous weapon system would have to be able to address it, though, as with distinction, reasonable judgments of proportionality would be highly dependent on the operational environment and battlefield in which the machine was deployed.[47] Again, assessing proportionality is one thing in close-in infantry urban warfare, but altogether different in remote uninhabited deserts, in the open ocean or undersea, or circumstances of machine-on-machine operations where few if any civilians are present.

Precautions in attack require that an attacking party take feasible precautions in the circumstances to spare the civilian population.[48] Precautions and feasibility, it bears stressing, are terms of art in the law of armed conflict that confer reasonable discretion on commanders undertaking at-

45. *See* WILLIAM H. BOOTHBY, THE LAW OF TARGETING 96–97 (2012) (doctrinal discussion of the difficulties of comparisons in specific cases versus general principles).

46. As Schmitt notes, human targeting officers in U.S. forces sometimes employ sophisticated software programming to estimate likely collateral damage from the employment of a particular munition in particular circumstances, given assumptions about a variety of factors. *See* Michael N. Schmitt, *Autonomous Weapon Systems and IHL: A Reply to the Critics*, HARVARD NATIONAL SECURITY JOURNAL 19 (Feb. 15, 2013), http://harvardnsj.org/2013/02/autonomous-weapon-systems-and-international-humanitarian-law-a-reply-to-the-critics/. These software tools are merely tools, dependent on the assumptions used and only address the likely collateral damage—not the comparison to military advantage. *See also* GEOFFREY S. CORN ET AL., THE LAW OF ARMED CONFLICT: AN OPERATIONAL APPROACH 194 (2012) ("One process for implementing ROE that is currently in use in Afghanistan by U.S. forces is the Collateral Damage Estimation Methodology or CDEM. . . . [T]he purpose of the CDEM is to cause the military to do a thorough proportionality analysis and then, depending on the estimated number of civilian casualties, seek approval for the attack.").

47. Thurnher, *supra* note 41 ("[F]or the use of an autonomous weapon system to be lawful, the system would be expected to reasonably distinguish between combatants and civilians . . . given the particular environment and circumstances of the battlefield ruling at the time.").

48. API, *supra* note 27, arts. 57, 58; for explanation, see, e.g., DINSTEIN, *supra* note 31, at 138–40; HENDERSON, *supra* note 39, at 157–96.

tacks.[49] The commander's obligation is grounded in reasonableness and good faith, and in "planning, deciding upon or executing attacks, the decision taken by the person responsible has to be judged on the basis of all information available to him at the relevant time, and not on the basis of hindsight."[50] Although some precautionary measures might have to be programmed into a genuinely and fully autonomous weapon system, in reality many of the precautions that might be at issue are not so much a question of what the weapon system does in a particular missile strike as part of a larger operation or battle, but instead are addressed in the planning for the overall attack as an operation, including the development of the rules of engagement. As such, and for the foreseeable future, the rules of precautions in attack fall upon commanders as they plan how to deploy their combat resources in an operation. Technology can be of great assistance in modeling likely harm to civilians or civilian objects or in deciding upon the munitions or weapon systems (including at some future point, perhaps, autonomous systems) to employ, but responsibility for the operation, including the weapons used, belongs to the commander.[51]

While it is true that for this and many other legal reasons a fully autonomous weapon system would likely have to be able to break off or alter an engagement—potentially even as it was being launched (in at least some circumstances, where "feasible")—just as a human commander might be required to do, precautions in attack operate generally at a level above an individual weapon, whether operated by a human soldier or by a computer integrated into it, and rely on pre-planning and information obtained beforehand. Situational awareness for those directly carrying out the attack is often very low, even if the target is visible to them, and, because they do not typically have awareness of the entire operation and its parts, they are often not in a position to make a judgment as to a change in precautions, and so cannot be held to account for a violation of reasonably planned precautions. The legal standard, in principle, would be no different for a

---

[49] We note that while the United States and Israeli armed forces exercise strict precautions in attack, their position on the customary legal status of such obligation is not publicly formalized.

50. *See* W. Hays Parks, *The Protection of Civilians from Air Warfare*, 27 ISRAEL YEARBOOK ON HUMAN RIGHTS 65, 110 (1997) (quoting German declaration on Article 57, Protocol I).

51. *See* Gregory S. McNeal, *Targeted Killing and Accountability*, 102 GEORGETOWN LAW JOURNAL 681, 739–50 (2014) (discussing the collateral damage estimation and mitigation process, and evaluating precautions in attack through weapon selection and use of weapons in targeting and targeted killing operations).

machine carrying out only one part of a larger operation, though policy might lead to greater caution. Precautions in attack, with respect to new weapon technologies, in the long run might turn out to be the most difficult requirement of targeting law for an autonomous system to meet, at least in circumstances in which there are civilians present,[52] precisely on account of the difficulties of anticipating and pre-planning conditions and, if the machine were to be genuinely on its own, identifying and executing attacks. But independence for such a machine system, divorced from the integration of all levels of planning for an attack and its execution, is frankly a long way away, if not entirely fanciful.

Many of the battlefield environments for which autonomous weapon systems are contemplated today do not involve such considerations, however; they are not situations with significant civilian presence. Still, consistent with other aspects of the interplay of automation technology and the law of armed conflict, incremental changes in technology will push for use of these systems in more and varied environments in which civilians are indeed present. Legal reviews will likely have to take into account many particularized facts—all the permutations of in- and on-the-loop interactions by soldiers with the machine, that, for example, is still short of full autonomy, but nonetheless has very large roles for automation.

Stepping back, it is critically important to understand that before an autonomous weapon system—like any weapon system—is used in a military operation, human commanders and those employing the weapon will generally continue to be expected to exercise judgment about the likely presence of civilians and the likelihood that they may be inadvertently harmed; expected military advantage; particular environmental features or conditions; the weapon's capabilities, limitations, and safety features; and many other factors. It is difficult to draw general conclusions in the abstract about the many complex legal issues involved in such scenarios. However, in many cases, even though a weapon system may be autonomous, much of the required legal analysis would be conducted by human decision makers who elect whether or not to use it in a specific situation. Whether legal requirements are satisfied in a given situation will still depend not simply on the machine's own programming and technical capabilities, but on human judgments as well. In thinking through these legal issues, it may be helpful also to distinguish between human decisions to deploy a system

52. *See* William H. Boothby, *Some Legal Challenges Posed by Remote Attack*, 94 INTERNATIONAL REVIEW OF THE RED CROSS 579 (2012).

(including decisions about settings for the system) in a particular circumstance and decisions made by the system itself once deployed. These two sets of decisions are linked in complex ways, and making the former decisions appropriately should reduce the possibility for error in the latter.

In sum there is no reason, in principle, why a highly automated or autonomous system could not satisfy the requirements of targeting law.[53] Like any otherwise lawful weapon, it depends on the use and the environment. Uninhabited deserts are different from urban warfare. Destroying rockets in the air—Iron Dome, for example—is different from rural counterinsurgency in which entering civilian villages is necessary, not just for bare security reasons, but to consolidate political relationships. In practical terms, autonomous systems might be better able to satisfy the law in some uses and environments than others, but that is not a matter of principle; it is a matter of whether and how far technological capability advances relative to the legal standard.

V. DEVELOPING AND CULTIVATING LEGAL RULES AND CODES OF CONDUCT

The conclusions that autonomous weapon systems do not inherently violate the longstanding and accepted rules of warfare, and that such systems can be effectively regulated under these rules, does not mean that this will be a simple or straightforward exercise. The advent of autonomous systems creates new challenges that need to be addressed.

It is quite rare for an international law-related question to arise before it actually becomes a real-life dilemma.[54] There is therefore a unique (although probably short-lived) opportunity to get it right; to develop the rules and code of conduct for such systems before they are fielded on the battlefields of the world in large numbers. Though primarily procedural, the best way to approach this and to effectively adapt the law of armed conflict to future autonomous weapons is a three-tiered approach.

At the highest level, some form of international instrument addressing autonomous systems is probably needed eventually. But the instrument that is needed is neither a blanket ban on autonomous weapons nor a moratorium on their development and deployment. At its most basic, such an

53. Schmitt & Thurnher, *supra* note 8, at 279–80.

54. Compare, for example, issues of autonomous weapon systems to cyber warfare issues, where the law is likewise struggling to keep up with technological and real-life developments.

instrument would make clear that autonomous systems are governed by the existing rules of warfare, that they are not inherently illegal thereunder and that they are subject to the customary law requirements of legal review of weapons.

Such an instrument could go much further, however, elaborating on two separate sets of substantive issues. First, it could include some form of interpretative application of the law of armed conflict to commanders deploying autonomous systems, explaining what information such commanders must have and what questions such commanders must ask before deciding to field the weapons in a given situation. Second, it could include rules and guidelines for the development of autonomous systems. Such rules and guidelines could be based not only on legal requirements, but also on policy considerations. For example, do we want to require all autonomous weapon systems to have a human override capability (some form of "kill switch")? Or perhaps we would like all such systems to have a built-in self-neutralization mechanism. Perhaps this instrument could include a minimum sensory requirement standard for autonomous systems that could be required to make reasonably-foreseeable distinction-based decisions on the battlefield?

This international instrument need not deal with all of these issues at once. Indeed, as explained below, a better approach is to reach consensus on some core minimum standards, but at the same time to retain some flexibility for international standards and requirements to evolve as technology evolves. Such an instrument is not likely to have compliance traction with States over time unless it largely codifies standards, practices, protocols and interpretations that States have converged upon over a period of actual development of systems. The process of this convergence will best be accomplished if it takes place gradually through informal discussions among States, informed by sufficiently transparent and open sharing of relevant information, rather than through formal treaty negotiations that if initiated too early tend to lock states into rigid political positions.

Such an international instrument could have many different forms. It could be a totally new legal instrument; it could be a new protocol under the CCW or it could be a more manual-based approach, similar to the recent *Tallinn Manual on the International Law Applicable to Cyber Warfare*,[55] but involving more direct participation and endorsement by States. Each ap-

55. TALLINN MANUAL ON THE INTERNATIONAL LAW APPLICABLE TO CYBER WARFARE (Michael N. Schmitt ed., 2013).

proach has its own advantages and disadvantages, but a variant on the *Tallinn Manual* approach may be especially appropriate here. That document was developed by an international group of legal experts commissioned by the NATO Cooperative Cyber Defence Centre of Excellence to develop and propose interpretive guidance for States' and other actors' consideration. Although the cyber context is different, insofar as there may be greater disagreement as to the appropriate legal framework, similar international processes—whether involving State representatives or independent experts or both—can help foster broad consensus or surface disagreements that require resolution with respect to autonomous weapon systems.

The second tier will be at the national level. All countries, and especially countries in which autonomous system development is currently ongoing, should promulgate their own national rules and policies addressing the two sets of substantive issues mentioned above. Undoubtedly many of the details of national rules and policies will need to remain secret, as they will involve sensitive matters of military capabilities and practices. States should be urged, however, to publish openly their general policies and to promote sharing of best practices. A step in this direction was the U.S. Defense Department's development and publication of a policy directive regulating automated and autonomous weapon systems that spells out limits and procedural requirements with regard to research, development and deployment of autonomous weapon systems.[56]

Finally, the third tier of a regulatory process should comprise the developers and potential users: the defense industry (including private sector developers of component technologies) and responsible militaries of the world. The former will need clear guidance as to what types of systems they should and should not be developing. The latter need to start thinking about methodologies, operating procedures, rules of engagement and other operational and doctrinal level rules for the use of autonomous systems. While both of these groups will be generally implementing standards and rules that flow down from the higher tiers, it is crucial for them to be involved in the discussions at the higher levels, as in many cases only they will have a clear understanding of the real technological and operational questions and dilemmas that arise. For example, collaboration between private sector weapon developers and the militaries that might acquire and deploy them will be critical in formulating and implementing effective test-

56. DoD Directive 3000.09, *supra* note 2.

ing and evaluation systems specifically for autonomous weapons.[57] It is preferable that all three tiers advance in parallel, each drawing upon the other in the complicated dialogue involving international and national interests that generally characterizes the development of international law.

Lawyers knowledgeable in the requirements of weapons legal reviews thus need to be engaged with weapon designers and engineers in the private sector from the beginning of the development process in order to help ensure that weapons law be as much a part of the system's design requirements as any other. Similarly, they must be engaged with military operators able to assist them in understanding the intended uses and battlefield environments for the system, as well as the capacities and limits of soldiers using the system and requirements for training of the operators. The development—not just the deployment and use—of a weapon system that is autonomous (or has an autonomous firing option) requires an "understanding of the legal parameters; the engineering design, production, and testing (or validation) methods; and the way in which the weapon might be employed on the battlefield."[58] While a review conducted during the acquisition process is not the final legal review, the stages of design, demonstration, manufacture and in-service deployment are "important stages for the input of formal legal advice."[59] To be meaningful, though, these legal parameters must translate into terms of reliability engineering that are "testable, quantifiable, measurable, and reasonable."[60]

Military lawyers are well aware of these issues and a robust debate over the legal review of increasingly automated weapons has been underway in internal defense ministry publications and journals dedicated to the law of armed conflict for well over a decade. One task is coordination and inte-

---

57. *See* REPORT OF ICRC EXPERT MEETING, *supra* note 7 ("[I]t is not clear how such weapons could be adequately tested given the absence of standard methods for testing and evaluating autonomous systems."). Backstrom and Henderson acknowledge the difficulties in testing systems, but add that "one pro-active step that could be taken as part of the legal review process is for lawyers to [provide] input [into] the test and evaluation phases by identifying areas of legal concern that could then be translated into testable elements." Alan Backstrom & Ian Henderson, *New Capabilities in Warfare: An Overview of Contemporary Technological Developments and the Associated Legal and Engineering Issues in Article 36 Weapons Reviews*, 94 INTERNATIONAL REVIEW OF THE RED CROSS, 483, 507 (2012).

58. Backstrom & Henderson, *supra* note 57, at 401.

59. Justin McClelland, *The Review of Weapons in Accordance with Article 36 of Additional Protocol I*, 85 INTERNATIONAL REVIEW OF THE RED CROSS 397, 401 (2003).

60. DEFENSE SCIENCE BOARD, U.S. DEPARTMENT OF DEFENSE, REPORT OF THE DEFENSE SCIENCE BOARD TASK FORCE ON DEVELOPMENTAL TEST & EVALUATION 38 (2008), *available at* www.acq.osd.mil/dsb/reports/ADA482504.pdf.

gration of review tasks across disciplines and between government defense agencies and the private sector, and to address questions about how to translate legal requirements into concrete engineering tests. A further coordination task is creating dialogue and cooperation among States and military establishments so that internally generated legal standards and requirements and engineering tests help develop informal norms, expectations, standards and shared understandings that might, at a suitable future point, crystallize into concrete national and/or international treaty rules with respect to highly automated or autonomous weapons.

The double challenge here is, on the one hand, for States to apply sufficiently clear and robust standards and rules, under reasonable understandings of their Article 36 obligations, as sophisticated, modern autonomous weapon systems are gradually fielded. And, on the other hand, ensure that standards and rules under Article 36 that States develop today will be equally relevant or adaptable for the future systems which will be developed ten, twenty, thirty years from now.

The fundamental principle underlying the gradual development of these standards and rules alongside the evolution of automation technologies, however, should be that what matters is ever greater compliance with the core obligations of the law of armed conflict: necessity, distinction, proportionality and humanity. Whether the actor on the battlefield is a "who" or a "what" is not truly the issue, but rather how well that actor performs according to the law of armed conflict. Debate over standards or rules for automated or autonomous systems should remain scrupulously neutral as between human or machine, and should affirmatively reject any *a priori* preference for human over machine. Even seemingly indisputable calls for a first principle of "meaningful human control" mistake the issue, which is lessening the harms of armed conflict within the law by the means that are the most effective. The principle of humanity is fundamental, but it refers, not to some idea that humans must operate weapons, but instead to the promotion of means or methods of warfare that best protect humanity within the lawful bounds of war, irrespective of whether the means to that end is human or machine or some combination of the two.

Moreover, as previously suggested, the solutions to the challenges posed by autonomous weapons likely lie in a process of gradual international legal development that evolves as technology advances. It is not one that seeks to predict how technology will or will not evolve, or one that tries to pause technological development until a new legal consensus is in place. Hitting the pause button on technology, pending consensus on the

legal framework to govern it, seems attractive for some of those who have concerns about the risks of such technologies. But there already *is* a general framework in place for these emerging weapons—the law of armed conflict and its fundamental principles, and the processes for the legal review of weapons.

VI. CONCLUSION

As increasingly automated—and in some cases fully autonomous—weapon systems enter the battlefield or become possible, it is important that international norms to regulate them head down a path that is coherent and practical. Contrary to the claims of some advocates, autonomous weapon systems are not inherently illegal or unethical. The technologies involved potentially hold considerable promise for making armed conflict more discriminating and causing less harm on the battlefield. They do pose great challenges, however, with regard to law of armed conflict rules regulating the use of weapons. Those challenges demand international attention and special processes for adapting existing law to meet those challenges.

Rather than seeking to impose, up front, a new set of prohibitory rules or seeking to suspend development of autonomous weapon systems pending a comprehensive agreement on rules to govern them, international regulation of autonomous weapon systems should begin with the premise that the law of armed conflict provides an appropriate general framework. States should work to build on that framework through continually-improving interpretive standards and agreed-upon best practices. There is no risk-free course of action, and the three-tiered approach we propose offers a way to appropriately and realistically constrain military activities, while improving adherence to core law of armed conflict principles.