13. Asleep at the switch? How killer robots become a force multiplier of military necessity

Ian Kerr and Katie Szilagyi

Given the tremendous military success of unmanned drones and other semi-autonomous weapons, a number of states are investing heavily in the development of the next phase of robotic warfare. Their hope is to develop fully autonomous weapons.[1] If deployed, these weapons – *killer robots*, as they have become known[2] – would be able to select their own targets and fire on them, without human intervention.

It is not difficult to see why some states are hard at work to develop killer robots while others seek to ban their use. In many ways, if they can be developed, killer robots would be *better* than human soldiers: "better, stronger, faster."[3] Killer robots would confer enormous military advantage to states with access to them, and a correspondingly destructive potential against the enemies of those states. For this reason, many people on either side of the argument see the development and deployment of killer robots as practically inevitable. Even those who recognize and fear the moral as well as physical dangers of killer robots feel compelled to consider their place in the next generation of battle. After all, "No nation wants to field people against faster, more accurate machines, and it is not entirely clear why they should have to."[4]

[1] These include the United States, Israel, Russia, and China. *See*: Human Rights Watch, *Shaking the Foundations: The Human Rights Implications of Killer Robots*. International Human Rights Clinic. Harvard Law School (2014), http://www.hrw.org/sites/default/files/reports/arms0514_ForUpload_0.pdf.

[2] Using such language may be a bit incendiary, but the phrase has grown in popularity as a descriptor for the very technologies we describe in this piece. See, *e.g.*, the Campaign to Stop Killer Robots, http://www.stopkillerrobots.org/.

[3] Kind of like the *Six Million Dollar Man* – minus the human being.

[4] These words are borrowed from Ryan Calo, offered as an important counter-consideration in his very thoughtful feedback on an earlier draft of this chapter.

There is a second central argument in favor of the deployment of killer robots. Substituting machines for humans in international conflict, we are told, not only will save lives by wielding the best big stick, but robots will also outperform human soldiers ethically.[5] Robots are not vulnerable to the perils that plague humans on the battlefield: exhaustion, elevated emotions, or the need to seek retribution for the death of a comrade. Advanced sensory capabilities also permit robots to cut through the fog of war – reducing confusion, friendly fire, and other erroneous responses. By virtue of their carefully programmed slave-morality, these sophisticated machines will be better able than humans to comport with international standards and the ethical rules of just war. With Asimovian aspiration, killer robots would be programmed, as Dr. Susan Calvin put it, to be "essentially decent"[6] – in a way that humans killers never could be. This, we are told, will reinforce and enhance international humanitarianism and reduce injustice in armed conflict.

Yet, Dr. Calvin's contention that robots are "worlds different"[7] than humans discounts not only the role that humans will inevitably have in programming and controlling them, but also the possibility that such robots, as Langdon Winner astutely warns, may have politics of their own.[8] By allowing these robots to exist, we ask, are we also allowing them to determine their own course? Are we allowing them to fundamentally transform the battlefield?

It is important to recognize that *both* of the central arguments in favor of killer robots are founded on a perception of their "inevitability." The first argument presumes that the development of killer robots is inevitable.[9] The second argument presumes that robot superiority is inevitable.[10]

Although it may *seem* inevitable that we should want to send robots rather than people into the line of fire, it is also true that allowing robots to make

---

[5] Ronald C. Arkin, Patrick Ulam & Alan R. Wagner, *Moral Decision-making in Autonomous Systems: Enforcement, Moral Emotions, Dignity, Trust, and Deception*, 100 Proceedings of the IEEE Special Issue on Interaction Dynamics at the Interface of Humans and Smart Machines 571 (2012).

[6] Isaac Asimov, I, Robot 178 (1950).

[7] *Id.*

[8] Langdon Winner, *Do Artifacts Have Politics?*, in The Whale and the Reactor: A Search for Limits in an Age of High Technology 19 (Langdon Winner ed., 1986).

[9] And, therefore, that the "good guys" had better develop and deploy them first.

[10] Already, unmanned systems are equipped with sophisticated sensors and processing power that promotes precision far better than humans can marshal on their own. *See* P.W. Singer, Wired for War: The Robotics Revolution and Conflict in the Twenty-First Century 863 (2009).

decisions about who shall live and who shall die (even if robots are functionally superior to humans) crosses a fundamental moral line. The deployment of killer robots would entail that we delegate crucial moral decisions of life and death away from robust human decision-makers in favor of relatively limited software algorithms. It would also make the decision to use lethal force easier, generating more rather than less armed conflict.

In this chapter, we investigate killer robots and their implications for international humanitarian law. In Section 1, we survey the current state of the art in military robotics, recognizing that although today's semi-autonomous technologies require a "human in the loop," it may not be long before human participation ceases to be a technological or military imperative. Section 2 considers the case for lethal autonomous robots and the technological project of programming ethical compliance with international humanitarian norms. To better understand the complexity of such demands, we offer in Section 3 a careful overview of the key considerations of international humanitarian law. In Section 4, we investigate the philosophical underpinnings of its purported approach – technological neutrality – and some of the problems inherent in that approach. In addition to its superficial and fictitious treatment of the technologies in question, we suggest that this approach permits a deterministic mode of thinking that expands the scope of that which is perceived of as "necessary" as a result of the adoption of the technology in question. In Section 5, we examine the implications of this in the international humanitarian law context, arguing that the "normative pull" of some emerging military technologies reshapes the rules regarding their permissible use. Consequently, we argue, even if killer robots can be said to comport with international humanitarian law, they will operate as a force multiplier of military necessity, thus skewing the proportionality metric and amplifying new forms of destructive, lethal force. We conclude in Section 6 by calling into question the appropriateness of international humanitarian law as the primary or exclusive means of regulating lethal autonomous military robots.

1. ROBOTIC WARFARE

It took less than a decade for robotic warfare to move from video game consoles[11] to the military theatre in the Middle East. Robotic weapons can be broadly classified into three categories: (1) *remote-controlled weapons* (e.g.,

[11] RoboWar is a freeware web-based game in which aspiring programmers design their own robot in the RoboWar-specific stack-oriented programming

unmanned aerial vehicles that require a human operator's confirmation before the weapon's launch sequence can be engaged);[12] (2) *semi-autonomous robots* (e.g., robot sentries that can accept sensory inputs and execute a specific action from a catalogue of set responses);[13] and (3) *lethal autonomous robots* (e.g., the yet-unrealized logical extension and oft-stated end goal of the current technology – machine systems capable of making tactical decisions and performing military operations independent of human input).[14]

Many of our most powerful weapon systems are already imbued with some degree of "autonomy" and, once initiated, are able to carry out operations independent of human control or oversight. For example, torpedoes[15] and cruise missiles[16] – precursors to today's Predator drones[17] – have long been able to determine (within a range of options) the optimal speed and trajectory required to engage a target without human presence and, sometimes, without human intervention.[18] A dramatic improvement over the fledgling application of air power in World War I, one example

---

language: RoboTalk. *See RoboWar*, SourceForge, http://robowar.sourceforge.net/RoboWar5/index.html (last visited Apr. 17, 2014).

[12] Patrick Lin, George Bekey & Keith Abney, Autonomous military robotics: risk, ethics, and design. Report for US Department of Navy, Office of Naval Research 51 (San Luis Obispo: Ethics+Emerging Sciences Group at California Polytechnic State University, 2007).

[13] Semi-autonomous robots have not departed completely from the standard of "human in the loop," but delegate increasing amounts of decision-making authority to the robot. *Id* at 105.

[14] *See* Ronald C. Arkin, Governing Lethal Behavior: Embedding Ethics in a Hybrid Deliberative/Reactive Robot Architecture. Technical Report GIT-GVU-07-11 4 (Atlanta: Georgia Institute of Technology, 2007).

[15] Torpedoes were one of the first technologies to self-select their targets, to some degree. U.S. mathematician John Von Neumann and a team of engineers designed a torpedo that received feedback from its environment – much like a household thermostat – to hone in on its destination. *See* Anthony D'Amato, *International Law as an Autopoietic System*, Developments of International Law in Treaty Making 10 (Rudiger Wolfrum and Volker Roben eds., 2005).

[16] Cruise missiles employ GPS technology to guide the missile to its target. The discovery of GPS was exalted more for its potential impact on weapons-guidance than for any other application. *See* C.B. Puckett, *In This Era of Smart Weapons, Is a State Under a Legal Obligation to Use Precision-Guided Technology in Armed Conflict?*, 18 Emory Int'l L Rev 645 (2004).

[17] Predator drones, a type of unmanned aerial vehicle, are one of the most well-known instances of robotic weapons. About 27 feet in length, Predators resemble "baby planes," although they do not have cockpits. In their first year of operation by the US military, Predators fired missiles at 115 targets in Afghanistan. *See* Singer, Wired, *supra* note 10.

[18] *See* Robert Sparrow, *Killer Robots*, 24 Journal of Applied Philosophy 62, 64 (2007).

of such "smart" bombs are precision-guided munitions, which employ "laser, electro-optical, or infrared guiding systems that keep them on course towards their targets."[19] These smart bombs emerged into widespread usage and the public consciousness after use by coalition forces in the 1990–1991 Persian Gulf War.[20] Their success at avoiding so-called collateral damage, particularly through preventing civilian casualties in targeted attacks in urban areas, was highly publicized.[21] By the time of the Gulf War, the circular error of probability of bombs dropped was a mere 10 feet, a dramatic improvement over the 3,300 feet expected during World War I.[22] Smart weapons have become standard for many military applications; indeed, some human-rights advocates even argue that only sufficiently smart weapons should be permitted to attack within urban areas.[23]

A significant number of currently operational military robots fall within the category of unmanned aerial vehicles. By its own assessment, the U.S. Department of Defense spent over $3 billion on these vehicles between 1990 and 2000, and another $4 billion from 2000 to 2010.[24] High-profile Predator drone attacks on suspected terrorists in Pakistan, Yemen, and elsewhere have resulted in numerous front-page headlines.[25] In 2009, the U.S. Air Force trained more remotely controlled aircraft pilots than actual fighter pilots.[26] New applications extend the already strong capabilities of the Predator. One such example, the Global Hawk, has been referred to as

---

[19] Danielle Infeld, Precision-Guided Munitions Demonstrated their Pinpoint Accuracy in Desert Storm; But Is a Country Obligated to Use Precision Technology to Minimize Collateral Civilian Injury and Damage, 26 GEO. WASH. J. INT'L L. & ECON. 109, 109 (1992).

[20] *Id*. at 109.

[21] *Id*. at 110.

[22] G.D. Bakshi, *Yugoslavia: Air Strikes Test of the Air War Doctrine*, 23 STRATEGIC ANALYSIS 791, (1999).

[23] C.B. Puckett, In This Era of Smart Weapons, Is a State Under a Legal Obligation to Use Precision-Guided Technology in Armed Conflict?, 18 EMORY INT'L L. REV. 645, 647 (2004).

[24] Sparrow, *supra* note 18 at 63.

[25] *See, e.g.*, Aliza Kassim, *Drone strike hits Pakistani tribal region*, CNN (Mar. 30, 2012), http://www.cnn.com/2012/03/30/world/asia/pakistan-drone-strike/index.html; *see also* Richard Norton-Taylor, *Rise of the Drones Poses Growing Dilemma for Military: MoD Confronts Moral and Legal Issues as Armed Robots Increasingly Take Warfare out of Human Control*, THE GUARDIAN (Apr. 2, 2012), http://www.guardian.co.uk/world/2012/apr/02/rise-of-the-drones-military-dilemma.

[26] Loes van Wifferen, Alienation from the Battlefield: Ethical Consideration concerning Remote Controlled Military Robotics (2011) (unpublished MA thesis, University of Utrecht) (on file at http://dspace.library.uu.nl/handle/1874/205856).

"the Predator's big brother."[27] It flies autonomously as opposed to being remotely piloted: an operator tells the UAV to take off; it then carries out its pre-programmed mission by acquiring directions in real time from its onboard GPS and operates independently until it returns and the pilot "hit[s] the land button."[28]

Land application of robotic weapons is more difficult and, as a result, less common. Robots have difficulty navigating uneven terrain. Nonetheless, the U.S. Army's Future Combat Systems project, currently in development, is aimed at creating a system for rapid deployment that would replace the current main battle tank with unmanned technology.[29] Early entrants to the realm of robotic land weapons include PackBot, a flagship product of the U.S. robot company iRobot.[30] First used in rescue efforts on September 11, 2001, to adeptly navigate Ground Zero, PackBot was later deployed to Afghanistan to act as a scout in treacherous cave systems.[31] In its latest deployment, a human makes the decision to fire the 5-millimeter light machine gun, but there is also an "automatic" mode where the robot can make its own decision.[32] South Korea aims to use this robot to shoot any human attempting to cross the demilitarized zone.[33]

Systems such as these continue to improve: components become smaller, computer processing becomes more powerful and less expensive, and weapons capabilities become more adept.[34] This Moore's Law-ish[35] trajectory of military robotics can be perceived in U.S. military operation in Iraq. When the forces first went into Iraq in 2003, only a handful of unmanned aerial vehicles were involved in the operation; by 2009, that number was 5,300.[36] At its outset, the operation had no ground robotic systems, but by 2009, over 12,000 such systems were in place.[37]

---

[27] Sparrow, *supra* note 18 at 92.

[28] *Id.*

[29] Sparrow, *supra* note 18 at 63.

[30] iRobot purchased its name from the seminal Isaac Asimov novel. *See generally supra* note 6.

[31] Singer, *supra* note 10 at 57.

[32] *Id.*

[33] *Id. See also*, Arkin *supra* note 14 at 5.

[34] *Id.*

[35] More than 40 years ago, Intel co-founder Gordon Moore forecasted the rapid pace of technology innovation. Moore observed that transistor density on integrated circuits had doubled about every two years from 1957 to 1965 and predicted that this would continue until at least 2020. *See* Gordon Moore, *Cramming More Components onto Integrated Circuits*, 38 Electronics (1965).

[36] P.W. Singer, *Military Robots and the Future of War*, Ted.Com (February 2009), http://www.ted.com/talks/pw_singer_on_robots_of_war.html.

[37] *Id.*

Both technological and military standards regarding such weapons are currently premised on a "human in the loop" – for now, human beings remain the gatekeepers of military decision-making.[38] However, if the aforementioned trajectory continues in a consistent fashion, it may not be long before human participation ceases to be a technological or military imperative. More than 40 countries are in the process of developing autonomous weapons.[39] Many of the systems under development will go far beyond "a 'fire and forget' system capable of determining its trajectory or pursuing its target to some limited extent."[40] We stand on the precipice of a military era that could ask us to decide whether we should delegate to machine systems the programming of missions, final targeting instructions – and even decisions about whether to pull the trigger or push the button.

2. THE CASE FOR LETHAL AUTONOMOUS ROBOTS

A key advantage of adopting lethal autonomous robots is that they could be programmed to circumvent common human frailties. Existing robots already outstrip their human counterparts in sensing their environments. With highly developed sensor systems operating at incredible speeds, these machines have advanced capabilities that allow simultaneous responses from many inputs.[41] This enables real-time analytical surveillance of the battlefield, thus reducing the "fog of war."[42] Robotic senses are not clouded by human emotions such as fear, hysteria, anger, or frustration.[43] They do not suffer from the human shortcoming of "scenario fulfillment"; the propensity to ignore or modify incoming information to align with pre-existing beliefs and ideas, and what Gary Marchant *et al.* describe as "a form of premature cognitive closure."[44]

Further, since robots need not emulate the human tendency for

---

[38] Lin, Bekey, and Abney *supra* note 12 at 70.

[39] Ugo Pagallo, *Robots of Just War: A Legal Perspective*, 24 PHIL. & TECH. 301, 315 (2011).

[40] Sparrow, "Killer" *Supra* note 18 at 64.

[41] Ronald Arkin, *The Case for Ethical Autonomy in Unmanned Systems*, 9 J. MIL. ETHICS 332, 334 (2010).

[42] *Id.* at 333. The "fog of war" refers to the uncertainty felt in a battlefield situation due to the volatile nature of that milieu and the unavailability of real-time updates.

[43] *Id.* at 334.

[44] Gary E. Marchant, et al., *International Governance of Autonomous Military Robots*, 12 COLUM. SCI. & TECH. L. REV. 272, 283 (2011).

self-sacrifice, they are better able to carry out operations conservatively. Similarly, such robots can be used in a self-sacrificing manner without triggering the guilt of a commanding officer or the need to steel oneself against human survival instincts.[45] Among other things, this could reduce the need for frontline human soldiers. As Peter Singer notes (tongue in cheek) about the "death" of a PackBot – a key member of most U.S. sentry teams operating in Iraq – when a robot dies, you do not need to write a letter to its mother.[46]

Will wars that involve more robots and fewer humans better conform to the ideals that undergird the international legal framework for just war? Or, will increasing automation and robotization of warfare be used, unjustifiably, to redefine what counts as "humane" according to international humanitarian standards?

3. THE NORMS OF INTERNATIONAL HUMANITARIAN LAW

The history of humanitarian law is inextricably tied to the development of new technology: as war technologies become increasingly advanced and capable of greater destruction, laws are put in place to limit that destruction.[47] These laws have been conceived in two distinct streams: (1) laws of general application that apply to all instances of warfare such that the imperative of humanity modulates how war is waged; and (2) specific international agreements that prohibit or limit the use of particular weapons, e.g., chemical and biological weapons.

One of first international efforts to proclaim standards for a humanitarian approach on the battlefield was the St. Petersburg Declaration of 1868.[48] The effort was convened to respond to a specific technological problem – the invention of bullets that would explode only on contact with a soft surface, such as the human body.[49] The St. Petersburg Declaration

---

[45] Arkin, "Ethical," *Supra* note 41 at 333.

[46] Singer, WIRED, *supra* note 10 at 52.

[47] Jakob Kellenburger, International Humanitarian Law and New Weapon Technologies, 34th Round Table on Current Issues of International Humanitarian Law, Keynote Address, ICRC (2011), http://www.icrc.org/eng/resources/documents/statement/new-weapon-technologies-statement-2011-09-08.htm.

[48] Arguably, there were a few other Conventions that predated the 1868 St. Petersburg Declaration, but this was the first one that issued a specific prohibition against a particular military tactic on the battlefield. *See also*, *Id.*

[49] Isabelle Daoust, Robin Coupland & Rikke Ishoey, *New Wars*, 84 INT'L REV. RED CROSS 345, (2002); *see* also the Declaration Renouncing the Use, in Time

banned the use of projectiles of less than 400 grams in conflicts.[50] It also decreed "the only legitimate object that states should seek to accomplish during war is to weaken the military forces of the enemy."[51]

With this standard in mind, the modern conception of humanitarian law began to emerge. International accords such as the 1899 Hague Declarations,[52] the 1907 Hague Conventions,[53] and the accompanying Regulations, provided a framework of general application that did away with any vestigial "anything goes" mentality for what was acceptable in combat. In particular, the inclusion of the Martens Clause, a compromise clause first included in the 1899 Hague Declarations, demonstrated the growing commitment to humanitarian precepts.[54] This clause stipulates that even in situations not governed by either customary or treaty law:

> . . . populations and belligerents remain under the protection and empire of the principles of international law, as they result from the usages established between civilized nations, from the laws of humanity and the requirements of the public conscience.[55]

This language continues to animate subsequent international law agreements. Concurrently, the stream of specific prohibitions began to expand as well. In 1925, the Geneva Protocol to the 1907 Hague Convention supplemented general prescriptions for appropriate behavior in battle

---

of War, of Explosive Projectiles Under 400 Grammes Weight, Saint Petersburg, 11 December 1868 (29 November by the Julian Calendar) LXIV *UKPP* 659 (1869).

[50] Kellenburger, "IHL," *supra* note 47.

[51] Daoust, Coupland & Ishoey, *New Wars*, *supra* note 49 at 346.

[52] *1899* Hague Declaration 3 Concerning Expanding Bullets, The Hague, 29 July 1899, UKTS 32 (1907), Cd. 3751 (Eng. Fr.).

[53] 1907 The Hague Convention for the Pacific Settlement of International Disputes, UKTS 6 (1971), Cmnd. 4575.

[54] The Martens Clause was a compromise clause first suggested by Fyodor Martens at the 1899 Hague Peace Conferences, which also appears in the 1907 Hague Conventions. It states that the "laws of humanity" undergird any adopted regulations, giving both combatants and noncombatants protections stemming from what would be expected per the civilized people and public conscience. This clause persists in the Additional Protocol I to the Geneva Conventions at Article 1(2), arguably prohibiting any weapons that are "abhorrent to the public conscience." *See* Daoust, Coupland & Ishoey, *New Wars*, *supra* note 49 at 351. *See generally*, Antonio Cassese, *The Martens Clause: Half a Loaf or Simply Pie in the Sky?*, 11 Eur. J. Int'l L. 187 (2000).

[55] Hague Convention (IV) Respecting the Laws and Customs of War on Land and Its Annex: Regulations Concerning the Laws and Customs of War on Land, The Hague, 18 October 1907, 2 Am. J. Int'l L. (1908) Supplement 117–27 (Eng. Fr.).

with a specific prohibition against chemical and biological weapons.[56] This marked the continued development of regulations against specific technologies.

In the immediate aftermath of World War II, global powers convened to establish new rules that would promote humanitarian principles on the battlefield. A need was identified to modify humanitarian law to deal with the exponential advancements in weapons technology.[57] The four Geneva Conventions of 1949,[58] direct descendants of the 1907 Hague Conventions, codified standards for humanitarian law. The treaties, which clarified states' responsibilities on the battlefield, were – at least in part – a response to the ever-advancing technology that was transforming the nature of war.[59] World War II introduced weapons of unprecedented lethality. The Geneva Conventions set out rules of general application to guide subsequent conflicts with humanitarian principles.

Yet, it was not only the weapons used to wage war that were transformed following World War II – the entire conflict paradigm shifted away from formal declarations of war. Accordingly, a contextual approach was required to better define when a situation could be viewed as an armed conflict. The subsequent 1977 Additional Protocols I and II to the Geneva Conventions,[60] which arose from an interest in protecting civilians, prescribed the responsibilities of parties to armed conflict within the international humanitarian law framework. Protocol I delineated expected behavior in international armed conflict, and Protocol II defined

---

[56] Protocol for the Prohibition of the Use of Asphyxiating, Poisonous or Other Gases and Bacteriological Methods of Warfare, Geneva, 17 June 1925, UKTS 24 (1930), Cmd. 3604 (Eng. Fr.).

[57] Thomas Michael McDonnell, *Cluster Bombs Over Kosovo: A Violation of International Law?*, 44 Ariz. L. Rev. 40, 64 (2002).

[58] Geneva Convention for the Amelioration of the Condition of the Wounded and Sick in Armed Forces in the Field, Aug. 12, 1949, 75 UNTS 31; Geneva Convention for the Amelioration of the Condition of Wounded, Sick and Shipwrecked Members of Armed Forces at Sea, Aug. 12, 1949, 75 UNTS 85, Geneva Convention Relative to the Treatment of Prisoners of War, Aug. 12, 1949, 75 UNTS 135; Geneva Convention Relative to the Protection of Civilian Persons in Time of War, Aug. 12, 1949, 75 UNTS 287.

[59] McDonnell, "Kosovo," *Supra* note 57 at 64.

[60] Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of International Armed Conflicts, June 8, 1977, 1125 UNTS 609; *see also*, Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of Non-International Armed Conflicts, June 8, 1977, 1125 UNTS 609.

how to appropriately use force in non-international armed conflict. Both protocols restrict military action by requiring uses of force to accord with humanitarian principles.

In this framework, the three principles are (1) proportionality, (2) discrimination, and (3) military necessity, which guide all tactical decisions. Adherence to these principles is designed to both minimize destruction during wartime and facilitate the peace-building process after a ceasefire is reached: the hope is that reconciliation is easier without malicious or indiscriminate attacks during the conflict. The critical legal question is the same irrespective of weapon used: does the specific use comply with international humanitarian law?[61]

3.1 Proportionality

Proportionality requires that any collateral injury to civilians and private property during a military operation must be balanced against the military advantage to be gained by carrying out that operation.[62] Quantifying a proportionate response requires balancing the repercussions of the action against the continuing threat that inspires the action.[63] The damage caused through the use of force in a military operation must, therefore, be proportionate to the strategic importance of using that force. Minor collateral damage can prevent an attack of minor consequence, while considerable collateral damage is permissible in situations when significant military advantage may be achieved. Attacks that cause damage exceeding the "concrete and direct military advantage anticipated"[64] from that use of force are disproportionate.

3.2 Discrimination

Discrimination requires an ability to distinguish between combatants and noncombatants during the course of a military operation. Civilians must not be directly targeted by a military operation.[65] Weapons and

---

[61] UNHRC, Report of the Special Rapporteur on extrajudicial, summary, or arbitrary executions, 14th Session Supplement No. 40, UN Doc A/HRC/14/24/Add.6 (2010) at ¶79.

[62] Protocol I, *supra* note 60 art. 57(2). *See also*, Enzo Cannizzaro, *Contextualizing Proportionality: Jus Ad Bellum and Jus In Bello in the Lebanese War*, 88 Int'l Rev. Red Cross 779, 785 (2006).

[63] Yoram Dinstein, War, Aggression and Self-Defence 17 (3rd ed. 2001).

[64] Protocol I, *supra* note 60 art 51(5)(b).

[65] Protocol I, *supra* note 60 art 51(1).

operational strategy must be selected to minimize collateral damage.[66] Yet, the principle of discrimination is complicated by the fact that civilians lose protection under this framework if they directly participate in hostilities.[67] Targeting combatants that are directly participating in hostilities has become an increasingly difficult task as the nature of warfare has transformed to involve nonstate actors as opposed to official forces.[68]

3.3 Military Necessity

Military necessity, as its name suggests, requires that an attack must be militarily necessary, i.e., the state must only defend itself or seek to guarantee its future security in executing the military operation – nothing more.[69] The principle is generally recognized to permit:

> only that degree and kind of force, not otherwise prohibited by the law of armed conflict, that is required in order to achieve the legitimate purpose of the conflict, namely the complete or partial submission of the enemy at the earliest possible moment with the minimum expenditure of life and resources.[70]

There are associated responsibilities to produce intelligence on the effects of the weapons used, the number of civilians that could be affected, and whether those civilians could take cover before an attack.[71] Attacks must have a concrete military objective; it is not permitted to simply spread terror among the civilian population.[72]

3.4 Humanity

The three integral principles of international humanitarian law are bolstered and informed by the complementary principle of humanity, which

---

[66] Protocol I, *supra* note 60 art 51(4). *See also* Dinstein *supra* note 63 at 119.

[67] Protocol I, *supra* note 60 art 51(3).

[68] Nils Melzer, *Interpretive Guidance on the Notion of Direct Participation in Hostilities Under International Humanitarian Law*, International Committee of the Red Cross (Jan. 7, 2009), http://www.icrc.org/eng/resources/documents/publication/p0990.htm at 79.

[69] Christopher Greenwood, *Historical Development and Legal Basis*, in The Handbook of Humanitarian Law in Armed Conflicts 36 (Dieter Fleck, ed., 1995).

[70] Melzer, *supra* note 68 at 79.

[71] Rapporteur, *supra* note 61 at ¶29; see also Michael N. Schmitt, *Military Necessity and Humanity in International Humanitarian Law: Preserving the Delicate Balance*, 50 Va. J. Int'l L. 795, 828 (2010).

[72] Protocol I, *supra* note 60 art 51(2).

is central to any legal analysis. The International Court of Justice confirms that "elementary considerations of humanity permeate international law."[73] Humanity inhabits a nexus among the three aforementioned principles. In particular, it is implicit within the principle of military necessity.[74] Proper respect of military necessity forbids the infliction of any suffering, injury, or destruction that is not required to accomplish the stated military objective.[75] This reduces the ambit of possible military action from everything not expressly prohibited, to only that which is truly required to achieve a legitimate military purpose.[76]

Other elements of this firm humanitarian basis for modern warfare are also codified in the *Geneva* framework. Several of these requirements are particularly informative for considering new technologies. Article 35 of Protocol I limits the right of the parties to choose weapons or means of warfare, prohibiting "weapons, projectiles and material and methods of warfare of a nature to cause superfluous injury or unnecessary suffering."[77] Interpreting this provision, the International Court of Justice has defined "a harm greater than that unavoidable to achieve legitimate military objectives." Also in Protocol I, Article 36 creates an obligation to assess whether a new weapons, means, or methods of warfare would be prohibited, either by the Protocol itself or other international law requirements.[78]

Notably, some states have developed mechanisms to consider the legality of weapons in accordance with Article 36.[79] Still, this remains an underutilized provision of the Geneva framework. Moreover, Article 57(2) requires that military forces must do "everything feasible" to verify the status and nature of the objective,[80] "take all feasible precautions" in planning and executing the attack,[81] "cancel or suspend" the attack if it

---

[73] *Corfu Channel (UK v Alb)*, 1949 ICJ 4, 22 (Apr. 9).

[74] Nils Melzer, *Keeping the Balance Between Military Necessity and Humanity: A Response to Four Critiques of the ICRC's Interpretive Guidance on the Notion of Direct Participation in Hostilities*, 42 INT'L L. & POL. 831, 908 (2010).

[75] M.N. Schmitt, T. McCormack & Louise Arimatsu, *Yearbook of International Humanitarian Law*, 2010 231 (2011). *See also* Melzer, *supra* note 68 at 79.

[76] *Id*.

[77] Protocol I, *supra* note 60 arts 35(1), 35(2).

[78] *Id*. art 36.

[79] These states include Australia, Norway, Sweden, the United States, Belgium, Canada, Denmark, Germany, and the Netherlands. *See* Daoust, Coupland & Ishoey, *New Wars*, *supra* note 49 at 354.

[80] Protocol I, *supra* note 60 art 57(2)(a)(i).

[81] *Id*. art 57(2)(a)(ii).

becomes apparent that the objective is improper or the loss of life would be excessive,[82] and issue "effective advance warning . . . unless circumstances do not permit."[83]

In addition, the principle of responsible command requires that an organization have a command structure capable of being aware of their obligations under the Protocol and of respecting humanitarian law.[84] Legal advisors must be available in armed conflict to advise on the applicable legal requirements.[85] These conditions demonstrate how humanity concerns are a cornerstone of operations conducted in accordance with humanitarian law.

3.5 Autonomous Weapons?

Let us imagine a new form of Predator drone that can operate independent of human interaction and is programmed to comply with humanitarian law. Without the participation of a military pilot with "eyes on target," would the autonomous drone be able to distinguish between a combatant and a civilian? What about a civilian carrying a gun for self-defense, who has never participated in the hostilities? If circumstances changed dramatically, could a robot reassess the situation and compose a proportional response? How would the robot infrastructure respond to a situation where what was computed to be militarily necessary in the situation was simultaneously disproportionate? Or indiscriminately targeted civilians? Would its computational commitment to the laws of war produce legal results? Or, would the robot subvert these principles to deliver a more efficient result? These are some of the questions that need to be answered if lethal autonomous weapons are to be permitted. But they are not the only questions.

4. TECHNOLOGICAL NEUTRALITY

As the above questions imply, the focal point in literature to date involves speculation about whether the next generation of military robots can, in fact, comport with the laws of war. Given the transformative impact of the behavior of states on international law, we contend that it is even more

---

82 *Id.* art 57(2)(b).
83 *Id.* art 57(2)(c).
84 *Id.* art 87.
85 *Id.* art 82.

important to consider how the introduction of lethal autonomous robots into the battlespace might impact international legal norms. We are not concerned with a robot breaking a particular law, but instead how robot action could (and will) subtly transform what states consider to be law. We are therefore less interested in how robots will adhere to the law as it currently stands than with how their participation in war might change international humanitarian law altogether.

To better comprehend this possibility, we begin with a suggestion. International humanitarian law can be understood as adopting a particular strategic framework for the regulation of emerging military technologies. This approach is known in other disciplines as the doctrine of *technological neutrality*.[86] Rather than implementing sector-specific rules or laws that are tailored to the functionality or capabilities of particular technologies, international humanitarian law rests on a set of foundational principles[87] that are said to be "neutral" with respect to any given technology. In this approach, military technologies are not regulated categorically or by class but through a determination of whether a particular implementation or use conflicts with the underlying principles of international humanitarian law.[88] Consequently, a military technology will only be limited or restricted if the manner in which it must be used or the results that it achieves is in conflict with international humanitarian principles.

The underlying approach has been successfully adopted in the regulation of other technologies. In the context of electronic commerce, technological neutrality:

> refers to statutory tests or guidelines that do not depend upon a specific development or state of technology, but rather are based on core principles that can be adapted to changing technologies. Since technological change is constant, standards created with specific technologies in mind are likely to become outdated as the technology changes.[89]

---

[86] Technological neutrality is a commitment to the idea that laws should be framed generally, as opposed to being designed for a specific technology. *See* Andrew Feenberg, Critical Theory of Technology 5 (1991), http://www.sfu.ca/~andrewf/books/critical_theory_of_technology.pdf.

[87] For example, the principles of distinction, proportionality, military necessity, humanity, etc.

[88] Conventions that prohibit or limit the use of specific weapons remain the exception: in most situations, there is insufficient political will for such agreements, even in the most self-evident cases, *e.g.* nuclear weapons. Conversations about outlawing weapons are inevitably tied to fears that those weapons will be used against you once you agree to get rid of your own.

[89] Michael A. Geist, *Is There A There There? Towards Greater Certainty for Internet Jurisdiction*, 16 Berkeley Tech. L. J. 1345, 1359 (2001).

Consequently, the standards adopted are deemed to be technology-neutral. The same standards can then be applied across a range of technologies. From a regulatory perspective, employing generic regulations is efficient, a way of avoiding modifications to the entire policy when the passage of time inevitably delivers new technologies.[90] In an effort to avoid re-inventing the legislative wheel for every emerging technology, the doctrine of technological neutrality tells us that we ought to guide our laws not by the technological state of the art but on the basis of sound legal judgments about the underlying functions that technologies aim toward. While policy-makers in electronic commerce law have enjoyed success in relying on this doctrine to date,[91] it is instructive to consider why technological neutrality might not be well suited for some emerging technologies.

Consider an example from the field of data protection (which also currently relies on technological neutrality as a guiding principle). Rather than making a separate privacy law for video rental records, another for surveillance cameras, yet another for facial recognition systems, and still another for social network sites, etc., the approach to data protection internationally has been to develop a core set of functional principles that are meant to adapt to various technologies that are likely to emerge over time. Consequently, most domestic privacy laws are derivative of a set of eight such principles, first articulated by the Organisation for Economic Co-operation and Development in the 1980s.[92] Focusing on fair information practices like collection limitation, purpose specification, use limitation, and the like, the privacy law's technology-neutral approach has been quite remarkable in its ability to regulate data collection, use, and disclosure across an array of technologies not previously anticipated when the data protection principles were themselves first enunciated.

Despite its success to date, technological neutrality is no panacea. More

---

[90] Ilse M. Van der Harr, Filomena Chirico & Pierre Larouche, *Network Neutrality in the EU, TILEC Discussion Paper No. 2007-030* (22–23 Sep. 2007), http://ssrn.com/abstract=1018326.

[91] See, *e.g.* UNCITRAL Model Law on Electronic Commerce with Guide to Enactment, United Nations Commission on International Trade (1996), http://www.uncitral.org/enindex.htm; and its Canadian counterpart, *The Uniform Electronic Commerce Act*, Uniform Law Conference of Canada (1999), http://www.law.ualberta.ca/alri/ulc/current/euecafin.htm.

[92] The ten principles of fair protection appended in Schedule I of Canada's Personal Information Protection and Electronic Documents Act (*PIPEDA*) are modeled after the privacy principles developed by the Organization for Economic Cooperation and Development (OECD). Personal Information Protection and Electronic Documents Act, S.C. 2000, c. 5. *See also*, OECD Privacy Principles, http://oecdprivacy.org/.

recently, disruptive technologies have begun to emerge that undermine or otherwise turn-on-its-head one (or more) of data protection's core regulatory principles. For example, ubicomp – a set of sensor networks and automating technologies devised to eliminate the need for human interaction during a series of information transactions[93] – disrupts the general data protection requirement of "knowledge and consent" in the collection, use, or disclosure of personal information. The practical impossibility of obtaining meaningful consent for every ubicomp transaction may result in the need for sector-specific regulatory approaches once society fully embraces the model of pervasive computing. As we suggest below, a similar need may arise in international humanitarian law. Put colloquially, some technologies change the game.

4.1 The Normative Pull of New Military Technologies

The introduction of a new military technology can reshape norms within military culture. Consider, for example, shifting standards in submarine warfare.[94] Humanitarian ideals have long informed the norms of naval warfare between surface vessels. Conflicts on the high seas were accompanied by standard responsibilities that aimed to preserve humanity in any altercations.[95] For example, military vessels were prohibited from attacking merchant ships and were instead required to capture and escort

---

[93] Ubiquitous computing, known colloquially among technophiles as "ubicomp," refers to a future where digital devices are integrated so seamlessly into our daily existence that no one notices their presence. *See* Mark Weiser, *The Computer for the 21st Century*, 94 Scientific American 94 (1991), http://www.ubiq.com/hypertext/weiser/SciAmDraft3.html.

[94] 1936 London Procès-Verbal Relating to the Rules of Submarine Warfare Set Forth in Part IV of the Treaty of London of 22 April 1930, London, 6 November 1936, UKTS 29 (1936), Cmd. 5302 (Eng. Fr.). Peter Asaro also uses it to exemplify problems associated with lethal autonomous robots; *see* Peter Asaro, *How Just Could A Robot War Be?*, Current Issues in Computing and Philosophy 50, 59 (Philip Brey, Adam Briggle & Katinka Waelbers eds., 2008). However, this example originates in Walzer's seminal work on just war. As a corollary to the "sink on sight" unlimited submarine warfare practices employed by German forces during World War II, German submarines stopped following the duty set out in the 1936 London Protocol to provide for the safety of the survivors of a sunken ship. This was justified under the auspices of military necessity: submarines are exposed to great danger if they have to surface and fulfill the obligations of surface vessels. The argument was made that the only alternative was to not use submarines at all or to use them ineffectively, which would have given control of the sea to the British navy, *see* Michael Walzer, Just and Unjust Wars 147 (1977).

[95] *Id.*

merchants to port. Likewise, after a naval battle, the successful vessels were required to rescue survivors by bringing them aboard.[96] The 1930 London Naval Treaty and its 1936 successor codified the law in this area.[97]

The historical example of the submarine illustrates the vulnerability of technological neutrality: as norms evolve and nations point to state practice to justify actions that stray away from – or are even in direct contravention of – international agreements, such conventions run the risk of becoming "blue law."[98] Customary international law may purport to fill this vacuum; yet, since international law depends on the development of norms based on the behavior of sovereign nation states, law can be transformed by a collective omission or new practice. And this can be achieved through the introduction of a new technology that "forces" new practices.

The 1936 London Naval Protocol reaffirmed that submarines had the same duties as surface vessels.[99] Consequently, submarines had a responsibility to comport with the longstanding obligations imposed on ships.[100] Attempting to follow these rules was not only impractical for submarines but also had the effect of imposing near impossible responsibilities. Lacking comparable crews to surface vessels, accompanying a merchant ship to port was not something submarines could feasibly do. Moreover, during World War II, German submarines were relegated to great depths for both their own safety and strategy since the Allied Forces controlled the surface.[101] Even if a submarine were to surface after a battle, the space constraints in a small cabin scarcely large enough for the existing crew and machinery meant that taking on two or three more people was out of the question.[102]

The most fundamental incompatibility between submarine operation and the constraints imposed on surface vessels went to the core of naval strategy: submarines were intended for stealth. They were used in situations where stealth was paramount. Surfacing would negate the defined military objective for which they were deployed. Before long, in the tumultuous context of World War II, any assumption that submarines

---

96 Jane Gilliland, *Submarines and Targets: Suggestions for New Codified Rules of Submarine Warfare*, 73 Georgetown L. J. 975, 981 (1984).

97 *Id.* at 978; *also see*, London Protocol, *supra* note 94.

98 Jane Gilliland describes a blue law as being one which is both clear on its terms and be clearly violated by an accused, but is unenforced because of changed conditions and long-term disregard by the community. *Id* at 989.

99 *Supra* note 96.

100 *Supra* note 96 at 989.

101 *Id.* at 981.

102 *Id.*

would comport with the Second London Naval Treaty disappeared. Accordingly, the treaty fell into disuse. New norms around the behavior of submarines emerged that were based on the way submarines were already being used in warfare. These norms were predicated – at least in part – on the way the technology had been designed.

Jane Gilliland cautions, "The law of armed conflict for submarines subrogates military necessity to humanitarian goals, and in so doing threatens the achievement of the humanitarian goals it seeks to protect."[103] The submarine example, which transpired in a world not yet constrained by international humanitarian law, showcases how the advent of new technology may sculpt international norms as what is easily practicable with the technology employed becomes the new widespread practice. Since international law is formed, to some degree, on the basis of state practice, this is troubling. The codifications that later occurred took into account the practices that had already unfolded on the battlefield. Failing to account for technological change thereby weakens the staying power of the codification of international law.

Of course, one could imagine a different and more humane historical outcome for the submarine. The fact that submarine vessels cannot easily rescue overboard combatants (who likely came to be that way because of a torpedo fired at them by the submarine) might just as easily be understood as a reason against their deployment rather than a reason in favor of excusing submarines from otherwise enforceable humanitarian obligations. In part, what history tells us is that the case in favor of military necessity is a strong one and that the technologies said to be necessary in carrying out important military objectives are not easily interfered with.

4.2 Technology: Intertwined with Ideology?

The doctrine of technological neutrality's basic assumption that we can, for particular legal purposes, treat all (or even most) technologies the same is further problematized when it is reduced to the more general proposition that technologies are themselves neutral. This form of technological neutrality treats technological tools as value-free, empty vessels, ready to fit the uses of their users.[104] "Guns don't kill people," goes the aphorism. "People kill people."

Although not frequently consulted in the literature on military technology and international humanitarian law, the field of science and

---

[103] *Id.* at 991.
[104] *Supra* note 86.

technology studies has had much to say about this form of technological neutrality. As Neil Postman so eloquently put it:

> Embedded in every technology there is a powerful idea, sometimes two or three powerful ideas. Like language itself, a technology predisposes us to favor and value certain perspectives and accomplishments and to subordinate others. Every technology has a philosophy, which is given expression in how the technology makes people use their minds, . . . in how it codifies the world, in which of our senses it amplifies, in which of our emotional and intellectual tendencies it disregards.[105]

Despite the compelling work of Langdon Winner and others who have sought to demonstrate that artifacts have politics,[106] we continue to "disregard the fact that many technologies determine their own use, their own effects, and even the kind of people who control them. We have not yet learned to think of technology as having ideology built into its very form."[107] To retell one of the best and most famous examples from the literature, let us consider the tomato harvester, a remarkable device developed by agricultural researchers at the University of California in the late 1940s.

Most people would ask – could something as straightforward as a mechanical tomato combine really have an ideology built into its very form?

The mechanical tomato combine cuts the plant from the ground, shakes the fruits loose, and – in newer models – sorts the fruits by size. In order to accommodate the shaking, scientists have bred new varieties of hardier, sturdier, and less tasty tomatoes. The machine replaced the handpicking system wherein human harvesters went through the fields multiple times to pick ripe tomatoes, leaving unripe fruits for later harvest. The machines reduce the cost of tomatoes by $5–7 per pound. The cost of these machines (~$50,000) means that only highly concentrated forms of tomato farming can afford this type of harvesting. Consequently, with the advent of the combine, the number of tomato farmers dropped from ~4,000 in the early 1960s to ~600 in 1973. An estimated 32,000 jobs were lost as a direct result of mechanization.

---

[105] Neil Postman, The End of Education: Redefining the Value of School 192–3 (1996).

[106] Langdon Winner, *Do Artifacts Have Politics?*, The Whale and the Reactor: A Search for Limits in an Age of High Technology 19 (Langdon Winner ed., 1986).

[107] Jerry Mander, Four Arguments for the Elimination of Television 350 (1978).

Winner contends that the broad adoption of the mechanical tomato harvester ultimately shifted the norms of tomato farming in California and, indeed, the nature of the tomato itself. His observations offer insight that can be applied to our considerations about the future adoption of lethal autonomous robots, and implications for the norms of international humanitarian law and the nature of war itself. But, before delving into that, let us first take a careful look at Winner's analysis by imagining a corollary set of norms that would seek to regulate farming by similar means.

We can imagine an environmental farming law that employs principles similar in nature to international humanitarian law. Eco-just farming might entail that those engaged in farming must likewise adhere, *mutatis mutandis*, to principles of: "proportionality," "distinction," and (let's call it) "agricultural necessity."

Proportionality, in this context, requires that any ecological harm caused by an agricultural operation must be balanced against the anticipated agricultural advantage gained by carrying out that operation. The principle of distinction, in this context, means that acts of farming should be directed only to agricultural products and not non-agricultural vegetation subsisting in its natural environment. The agricultural necessity principle, in this context, means that an agricultural intervention must be intended to help achieve an agricultural objective, and the resulting ecological harm must be proportional and not excessive in relation to the concrete and direct agricultural advantage achieved through the intervention.

We can imagine entrenching these norms as a means of safeguarding the environment against the potential evils of modern agriculture. And yet it is not difficult to see that the outcomes described by Winner would be reproduced. If we can assume that the means by which the harvester is employed does not implicate other non-agricultural vegetation in its natural environment, the adoption and use of the mechanical tomato harvester would easily be justified in terms of its agricultural necessity and said proportionality.

The necessity requirement would be achieved simply by ensuring that whatever is deemed an agricultural necessity is characterized as sufficiently important to trump any resultant harms. In this case, the agricultural need to feed many people in and outside of California is met through the enormous increase in yield, the significant reduction in cost, and the incredible overall efficiency in production. Meeting agriculture objectives in this way would be understood as a social benefit that eclipses any sacrifices to the marginal practice of rural agricultural culture. Applying the Doctrine of Double Effect, since the foreseeable harm to rural agricultural practice was inextricably intertwined with the agricultural objectives of increased productivity, efficiency, and cost-effectiveness in the harvesting

process, and since the introduction of the mechanical tomato harvester was primarily intended to achieve these overall goods, its adoption will be justified.

The proportionality calculus would generate a similar outcome. Recall that proportionality requires that an agricultural intervention seeking to fulfill an agricultural objective must not be undertaken if the ecological harm is known to outweigh the anticipated agricultural advantage. As is often the case, the harms cannot be measured until the technology is in place. Of course, once it is in place, we are no longer talking about some neutral cost-benefit analysis. The very adoption of this cumbersome and expensive machine system, the layoff of crew upon crew of farm workers, the reconfiguration of farmland, and so on, are themselves an assertion of a political preference for one set of values (productivity, efficiency, cost-effectiveness) over another (homegrown, authentic, domestic). The introduction of the mechanical tomato harvester is an assertion of one way of life and the negation of another. Through a sleepy haze, we almost need to remind ourselves: what was the original agricultural objective in introducing the mechanical tomato harvester?

Upon wakeful reflection, the objective was to *better harvest* tomatoes, which, it turns out, is a completely different goal than to harvest *better tomatoes*. Recall that the only shortcoming of this otherwise incredible agricultural device was that it could not handle soft, juicy (delicious) tomatoes. As Winner described it, the introduction of this technology necessitated plant breeders to come up with "new varieties of tomatoes that are hardier, sturdier, and less tasty than those previously grown."[108]

Asleep at the switch, we practically forget that by permitting this technology, we also permit it to determine its own use.

Perhaps it is because of this propensity toward technological somnambulism that, during and after a carefully crafted proportionality analysis, we tend to be dozy in recognizing that the entire balancing act was essentially dictated by the characterization of the objectives and the perceived "necessities" of the situation. "We *need* a new form of tomato," the engineer later tells us. Of course, this was not originally an agricultural necessity. But it became perceived as such just as soon as the social investment in mechanical tomato harvesters was made. Technology can shape our perceived needs. So much the worse for tasty tomatoes.

It is in this sense that technology is not neutral and can be used (if we let it) to reshape social norms. As Winner was so clear to point out, this

---

[108] *Id*. at 22.

observation does not entail technological determinism.[109] Rather, it recognizes that what technology makes possible has the power to generate in our minds what may later be perceived of as necessary. But only as a result of the adopted technology, which subsequently permits (but does not necessitate) the cart to drive the horse. The so-called agricultural necessity for a new tomato is not solely the consequence of adopting the mechanical tomato harvester. The ability to see the new tomato as an agricultural necessity is only perfected when plant breeders actually invent a technique for creating one.

At the same time, technology can also induce a related form of dogmatic slumber that permits another crucial fallacy to occur – an illogical inversion of Kant's famous insight that "ought implies can."[110] Through a strange form of grammatical alchemy we mistakenly come to believe that because the technology makes something possible, it also makes it necessary. Our perceived needs are thereby (re)shaped by our sense of what is possible. This propensity is crucial to remember in the application of balancing provisions. How technology shapes our perceptions will have a significant impact on our understanding of what is proportional and the means by which we apply principles like agricultural necessity.

Drawing upon Winner's example, outcomes in the military context will further assist in demonstrating the limits of technological neutrality. In the next section, we examine the interaction between lethal autonomous robots and the norms of international humanitarian law. First, we briefly investigate the possibility of robots comporting with humanitarian law. Second, we consider the normative pull of military technologies, arguing that the mere introduction of some military technologies can actually alter prior norms and practices. Finally, we attempt to show how shifting battle norms might occur through the introduction of lethal autonomous robots. We argue that robotic technologies act as a force multiplier in the

---

[109] Langdon Winner, *Technologies as Forms of Life*, PHILOSOPHY OF TECHNOLOGY 107 (David Kaplan ed., 2004). *See generally* Robert Heilbroner, *Do Machines Make History?*, 2 TECH. & CULTURE 335, (1961); *see also* William H. Shaw, *The Handmill Gives You the Feudal Lord: Marx's Technological Determinism*, 18 HISTORY AND THEORY 155, (1979); *see also* Alvin Hansen, *The Technological Interpretation of History*, 36 Q. J. OF ECON. 72 (1921).

[110] The maxim "ought implies can" is a form of transcendental idealism that leaves open the possibility that we have free will, *see* IMMANUEL KANT, RELIGION WITHIN THE BOUNDARIES OF MERE REASON 6:50. In the cyberspace context, Lawrence Lessig cautions against the use of sweeping rhetoric about the nature of technology, since it can lead to deterministic conclusions, *see* LAWRENCE LESSIG, CODE: VERSION 2.0 31 (2006).

determination of military necessity, amplifying the amount of permissible destructive force in carrying out an operation.

5. LETHAL AUTONOMOUS ROBOTS AND INTERNATIONAL HUMANITARIAN LAW

5.1 Comporting with International Humanitarian Law

Many proponents believe that lethal autonomous robots will one day reach the level of sophistication necessary to comport with international humanitarian law. The threshold question for the deployment of such a machine is whether it would be capable of selecting appropriately specific targets to achieve the standard required by the discrimination principle. In light of its autonomous operation, discrimination is an important primary consideration since it is a binary standard. A weapon that cannot distinguish between combatants and noncombatants is indiscriminate – and therefore *prima facie* illegal – as it is fundamentally incompatible with the laws of war.[111] One strategic response to an inability to meet the discrimination norm is to limit the use of robots to targeting weapons rather than people.[112] However, even with this strategy, adverse results are foreseeable. How might a machine system differentiate between a friendly peace officer carrying a service weapon and a guerilla warrior with a similar gun? Demarcating the distinction between civilians and combatants is further complicated in context-dependent situations, such as a combatant wishing to surrender or an enumerated military target no longer posing sufficient threat to constitute a legitimate target.

Even if this threshold question were to be adequately addressed, others remain. Ugo Pagallo identifies five factors that would need to be identified before a robot could legally engage a target. These factors, each deriving from a precept of international humanitarian law, are enumerated as follows:

(1) responsibility of humans who grant use of autonomous lethal force;
(2) military necessity in fixing criteria for the target;
(3) discrimination of the target identified as a legitimate combatant;

---

[111] Protocol I, *supra* note 60 art 51(4)(c); *see also* Marchant, *supra* note 44.

[112] *See* John Canning, *A Concept of Operations for Armed Autonomous Systems*, Speech at 3rd Annual Disruptive Technology Conferences (Sept. 6–7, 2010), http://www.dtic.mil/ndia/2006disruptive_tech/canning.pdf.

(4) principle of double intention so as to define tactics for engagement, approach and stand-off distance; and
(5) proportionality in selecting weapon firing patterns.[113]

A mere glimpse at the overview of international humanitarian law set out above in Section 3 reveals that the intellectual and sensory processes required to engage and operationalize each of these factors are multi-faceted and richly layered. The balancing functions required by these complex norms are difficult to reconcile with the Boolean logic and other current argument schemes that robots employ in order to render decisions. How would a robot properly assess the importance of a military target? In an international humanitarian law framework prescribing the assessment of proportionality as the intervening quality between principles of military necessity and humanity, the bar is set quite high, requiring the artificial intellect to assign three separate values and then correlate between three highly subjective variables. As Lin, Bekey, and Abney astutely identify, this framework is considerably more complex than the simplistic hierarchy of Asimov's laws.[114] And, as the frequent subversion of the robotic laws in Asimov's short stories reveals, the resolution of perceived conflicts between the norms within the framework can produce significant unanticipated results. Proportionality and military necessity are even more sophisticated, context-dependent concepts. Moreover, this does not account for challenges such as the emergent behaviors described by Ray Kurzweil: machine learning can generate unanticipated results.[115]

Kenneth Anderson describes the problem robots would encounter when assessing proportionality for *jus in bello* as a comparison of apples and oranges – an exercise in weighing incommensurate factors.[116] A decision heuristic based on simple inference would be inadequate for making complicated battlefield decisions; instead, robots would need to learn by doing and develop appropriate skills through practice.[117]

---

[113] Pagallo, "Just War," *supra* note 39 at 11.

[114] Lin, Bekey, and Abney *supra* note 12 at 76. Asimov robot laws first appear in his short story "Runaround," part of his seminal anthology *I, Robot*, *supra* note 6.

[115] Emergent behavior refers to the complexity that results when intelligence becomes self-organizing. Futurist Ray Kurzweil anticipates the same phenomenon will result when computers exceed human intelligence. Ray Kurzweil, The Age of Spiritual Machines (1999).

[116] Kenneth Anderson, *The Ethics of Robot Soldiers*? Law of War and Just War Theory Blog (Jul. 4, 2007), http://kennethandersonlawofwar.blogspot.com/2007/07/ethics-of-robot-soldiers.html.

[117] *Id.*

Even if roboticists are able to overcome many of the above challenges, programming robots to simply accord with international humanitarian principles may prove insufficient. Many nations that engage in warfare see these precepts as a minimum threshold; however, moral behavior is usually thought to significantly surpass minimum standards. After all, mere adherence to the letter of international legal conventions has the potential for catastrophic results. A rote application of bare minimum principles could justify significant loss of human life as collateral damage, while a more carefully crafted operation (or choosing not to undertake operations in light of the risks) could avoid that loss of life altogether.

Some proponents of lethal autonomous robots remain deeply convinced that it would be possible to design ethical programming that would not merely conform to, but ultimately surpass, the normative requirements of international humanitarian law. Arkin, for example, contends that case-based reasoning, already employed in other intelligent robotic systems, could prove useful in this regard.[118] He believes that such military applications would eclipse the performance of soldiers on the battlefield: "I am convinced that they can perform more ethically than human soldiers are capable of."[119] The fact that robots might sometimes fall short of the standard of the laws of war does not mean that they are worse than humans. Humans also fall short of this standard with what Arkin characterizes as "depressing regularity."[120] Using comprehensive system mapping and logical expressions, he describes the proposed functionality of lethal autonomous robots, including "architectural design options," that would inject moral reasoning capability into robots.[121]

It remains unclear how realistic the technological project of programming ethical compliance is in such a complex area of law. However, in our view, there are additional matters to consider.

5.2 Lethal Autonomous Robots as a Force Multiplier of Military Necessity

Both the submarine and the mechanical tomato harvester were game-changers in terms of what was subsequently seen as necessary and proportional in naval battle and tomato agriculture. The basic mechanism in each case remains the same: when a disruptive technology changes the nature

---

118 Arkin, "Governing," *supra* note 14 at 12.
119 *Id.* at 7.
120 Arkin, "Ethical," *supra* note 41 at 1.
121 *Id.* at 61.

of what is possible, it also expands the scope of inclusion for what can prospectively be perceived of as "necessary."[122] The consequences of this for international law are significant. If norms can be shifted in a manner mandated by the technology, then its potential to transform international law – *where practice becomes principle* – is enormous.

The power to induce a shift in norms in this way has led many academics to register concern about technology's influence over international law. Colin Picker expresses his worry as follows:

> Perhaps most problematic is the [fact] that technology is a determinate force that acts as an invisible hand creating, shaping and destroying international law. Failure to handle such a powerful force will result in policy makers essentially abdicating the international regime to technology.[123]

With many references peppered throughout his chapter to "the invisible hand of technology,"[124] Picker seems to be suggesting that the appropriate underlying philosophical worldview for international law is technological determinism – the idea that technology determines social outcomes.[125] Picker later explicitly denies this, claiming:

> I am not arguing, however, in favor of technological determinism. Technological determinism implies a stronger and more comprehensive relationship between technology and international law than I would assert exists. Policy makers can ignore technology, but at a tremendous cost.[126]

Our position is somewhat different. To us, it is not as though the invisible hand of technology magically removes all other social outcomes or

---

[122] Better prediction technology makes possible an argument that preemption is necessary. For instance, the much maligned "Bush doctrine" justifies pre-emptive self-defense in the context of the "War on Terror" through advanced technological prediction capability that better informs government agencies about the threat of a terrorist attack. *See United States Military Academy, President Bush Delivers Graduation Speech at West Point*, The White House (Jun. 1, 2002), http://georgewbush-whitehouse.archives.gov/news/releases/2002/06/20020601-3.html.

[123] Colin Picker, *A View from 40,000 Feet: International Law and the Invisible Hand of Technology*, 23 Cardozo L. Rev. 149, 151 (2001).

[124] *Id.*

[125] The phrase "technological determinism" was reportedly coined by economic industrialist Thorstein Veblen. *See*, Thorstein Veblen, Engineers and the Price System 38 (2001). For a fulsome survey of philosophers with varying view on technological determinism, *see* Bruce Bimber, *Karl Marx and the Three Faces of Technological Determinism*, 20 Social Studies of Science 333 (1990).

[126] Picker, "View," *supra* note 123 at 203.

possibilities, or that "it" somehow punishes those who do not respond accordingly with heavy costs. Rejecting this strange metonym, we believe that a more telling metaphor explaining the nexus between technology and norms is the one ultimately adopted by Langdon Winner:

> A more revealing notion, in my view, is that of technological somnambulism. For the interesting puzzle in our times is that we so willingly sleepwalk through the process of reconstituting the conditions of human existence.[127]

It is not that technology actually compels us to adopt certain norms (or that it actually puts us into a sleep-like trance). Rather, our social uses of technology can reconstitute our preconceptions so that we will not easily experience other existing possibilities.[128] "Of course we need new breeds of [tasteless] tomatoes – how else would we efficiently harvest them?" "Of course we cannot require submarine pilots to save overboard enemy combatants [whose boats they have blown apart] – how otherwise would they effectively carry out their missions?"

It is perhaps trite to say that international law was intentionally constructed to provide an extremely flexible framework.[129] Picker's more interesting claim is that a primary reason for doing so is to accommodate the protean nature of technology. Like the more recent technology-neutral frameworks used in electronic commerce and data protection law (both of which are derived domestically from international models), the core design of international humanitarian law is consistent with promoting, rather than restricting, innovation. While the technology in question is itself conceived of as neutral, the framework said to regulate it is not: it is designed to encourage and accommodate the overall use of technology.

When value-neutral approaches are applied to deeply value-laden technologies, the results can be disingenuous. Recall that the four Geneva Conventions underlying international humanitarian law were concluded in 1949. Animating this process was not only the aftermath of World War II, but the specter of weaponization in the newly arrived nuclear age. The process was concluded on August 12, 1949, less than a month after the

---

[127] Winner, "Forms of Life," *supra* note 109 at 107.

[128] Neil Postman, Technopoly: The Surrender of Culture to Technology (1993).

[129] The multiple sources of international law in the ICJ Statute is not intended as a single behavioral code, but instead as a flexible entity that recognizes a pluralistic legal perspective. *See* Statute of the International Court of Justice, June 26, 1945, 59 Stat 1055, art 38(1).

USSR test detonated its first nuclear weapon.[130] Yet, efforts to categorically prohibit such weapons of mass destruction were met with resistance. Unlike previous consensus over combat gases and biological weapons, the international community was unable to establish the political will to prohibit the development of nuclear arsenals.[131]

It was not surprising – in the golden age of technological neutrality – to see that very approach adopted by the International Court of Justice (ICJ) in its Advisory Opinion on the Legality of Nuclear Weapons.[132] Unsatisfied with the international community's failure to clarify the legality of nuclear weapons, the World Health Organization had asked the ICJ to adjudge on the legality of using nuclear weapons. The ICJ was resistant to the idea that one technology could be expressly forbidden and instead clung to the precepts of international law: the weapon itself was not illegal; the acts that could be committed with that weapon were what was illegal.[133] And provided that the weapon was used in a manner that satisfied the legal test for proportionality, discrimination, and military necessity, there was no need to outlaw it outright. The Court discarded arguments put forward by states that suggested that "nuclear weapons would be illegal in any circumstances owing to their inherent and total incompatibility with the law applicable in armed conflict,"[134] instead stating that:

> the principles and rules of law applicable in armed conflict – at the heart of which is the overriding consideration of humanity – make the conduct of armed hostilities subject to a number of strict requirements. Thus, methods and means of warfare, which would preclude any distinction between civilian and military targets, or which would result in unnecessary suffering to combatants, are prohibited. In view of the unique characteristics of nuclear weapons . . . the use of such weapons seems scarcely reconcilable with respect for such requirements. Nevertheless, the Court considers that it does not have sufficient elements to enable it to conclude with certainty that the use of nuclear weapons would necessarily be at variance with the principles and rules of law applicable in armed conflict in any circumstance.[135]

Despite the inherent flexibility of international humanitarian law, it is still difficult to imagine circumstances in which the use of a nuclear

---

[130] Kim Gordon-Bates, *The ICRC and the Nuclear Weapon: the Story of an Uncomfortable Paradox*, ICRC (Mar. 18, 2003), http://www.icrc.org/eng/resources/documents/misc/5krbdw.htm.

[131] *Id.*

[132] International Court of Justice, Advisory Opinion Report 226: Legality of the Threat or Use of Nuclear Weapons (1996).

[133] *Id.* at ¶¶74–87.

[134] *Id.* at ¶95.

[135] *Id.*

weapon could satisfy its norms. Nuclear weapons will, in their present form, be consistently unable to discriminate between civilians and combatants. Accordingly, any use of a nuclear weapon that satisfies the criterion of distinction would be an operation calculated to annihilate an entire area. As we have seen, such an act is clearly inhumane and almost certainly disproportionate to any act it purportedly answers.[136] Moreover, and perhaps even more troubling, allowing a nuclear weapon to remain within the *arsenal of possibility* might permit its use to be justified prospectively under the guise of military necessity in subsequent situations. As Justice Higgins describes in the Nuclear Weapons decision, questions of numbers of casualties or inflicted suffering "must be resolved as part of the balancing equation between the necessities of war and the requirements of humanity."[137]

With all of this, we see that the framework for balancing international humanitarian norms is sufficiently malleable to permit destruction and lethal force. While it is true that any military action must be constrained within the parameters of a proportional response, the overall potential for destruction is unquestionably augmented by the existence of certain destructive and lethal technologies with advanced capabilities.

As an illustration, let us imagine these norm-conflicts along a continuum. At one end of the continuum are outcomes premised solely on humane or humanitarian grounds (the principle of humanity). At the other end are outcomes that focus exclusively on carrying out destructive or lethal military objectives (the principle of military necessity). Somewhere in the middle, where these two norms are in direct conflict, the adoption of a new technology is often sought as a military solution.[138]

But the introduction of such a technology is often (to use a military metaphor) a "force multiplier"[139] of military necessity. As we suggested in our Kosovo example and with the subsequent questions raised by lethal autonomous robots, the ability to capitalize on military possibilities created by such technologies raises the stakes in terms of possible military objectives and operations that were unfathomable prior to the emergence

---

[136] Conceivably, one might argue that it would be proportionate to use a nuclear weapon in response to another nuclear attack.

[137] *Id.*, *see* dissenting opinion of Judge Higgins at ¶¶14, 20.

[138] The case for lethal autonomous robots set out above in Section 2 provides an excellent example of this.

[139] The term "force multiplier" refers to a factor that significantly enhances the effectiveness or strategic advantage of a particular force. Arkin expects lethal autonomous robots would result in force multiplication by having robots and soldiers side-by-side on the battlefield. *See*, Arkin in *supra* note 5 at 13.

of the technology. This allows us to "sleepwalk" toward a perspective that sees various uses of these technologies as a military necessary for resolving present and future armed conflict, even if they result in more death and destruction.

There is an interesting connection here between the concept of military necessity and the worldview of technological determinism that we reject. If one listens closely to the justifications often given in support of the military necessity in the use of a particular technology, there is a false strand of determinism running through it. The ICJ commits this very fallacy in contemplation of a situation wherein a state's "very survival would be at stake," anticipating that nuclear weapons may be the only recourse in "an extreme circumstance of self-defence."[140] It is this deterministic thread built into the very fabric of military necessity that makes it a force multiplier.

How might this show itself in the case of lethal autonomous robots?

In its most utopian form, robotic warfare seeks to remove humans from the battlespace, anticipating fewer people fighting and fewer casualties. Of course, in almost all conceivable situations, casualties would remain – it is only friendly casualties that are reduced by replacing one side's soldiers with robots. Failing to acknowledge this reality risks callousness in attack. Either way, the existence of lethal autonomous robots will surely be a force multiplier of military necessity in terms of the general military objective of reducing friendly casualties. "If we have expendable mechanical mercenaries that we can send into battle instead of our children, how could we *not* do so?"

Once we do so, the use of lethal autonomous robots will have a profound effect of lowering the threshold for entry into war: war will be (domestically) perceived of as easier both politically and logistically if there are fewer people involved. Fewer soldiers need to be recruited and fewer deaths have to be justified to win public support of the war effort. As Pagallo astutely identifies, autonomous robots have a double effect: they impact the traditional *jus in bello* rules for a fair fight on the battlefield while the capacity of the technology simultaneously influences a nation's political decision to go to war.[141] Again, one can view this as a force multiplier of military necessity. "If we have robots on the ground that can carry out an important military operation with few or no friendly casualties, why wouldn't we engage the enemy on this mission?"

To the extent that lethal autonomous robots will still interact with

---

[140] *Supra* note 132 at ¶96.
[141] Pagallo, *supra* note 39 at 303.

human military collaborators, the force multiplier of remote-controlled warfare can also have individual psychological effects on the soldiers engaged in it, increasing sympathy for military objectives and military necessity by minimizing our empathy regarding circumstances no longer seen through the lens of humanitarian ideals.[142] Today's warriors often fight from some cubicle, operating aircraft remotely, and developing a "Playstation mentality" with regard to the waging of war.[143] YouTube footage of Predator drone attacks set to music and shared and celebrated among soldiers online further distances the acts from the actors.[144] Increasing both the physical and psychological distance between soldiers and their targets not only dampens respect for human life – it also makes it easier to follow military objectives, especially those perceived of as necessary. After all, those being killed are only ever encountered as pixels on a screen.

This asymmetric element of modern warfare is not only dangerous; it also conceptually challenges the foundations of war by skewing the balance between humanitarian ideals and military necessity. How does the very nature of conflict change if one side's soldiers are never actually in danger? It is already the case that an American soldier can serve an entire tour of duty in Afghanistan or Iraq, work 16-hour days, and still eat lunch every day at a Carl's Jr. restaurant just outside the gates of his or her Nevada Air Force base.[145] While it may be appealing to imagine being at war with an enemy without experiencing casualties, it also solicits a larger question: If one side has no people in harm's way, is it truly a war?[146] And, more to the point for present purposes, without a significant level of human investment, will that side be able to see, understand, and inculcate the humanitarian norms in tension during battle? The advent of lethal autonomous robots has the potential to greatly exacerbate these risks in a manner that could be profoundly destabilizing to the framework of armed conflict.

---

[142] *See e.g.* Singer, WIRED, *supra* note 10; Asaro, *supra* note 94; Sparrow, "Killer," *supra* note 18; van Wifferen, *supra* note 26.

[143] Van Wifferen, *supra* note 26 at 38.

[144] These videoclips, many of which are freely available on YouTube, are colloquially known as "war porn." Peter Singer cites a particularly egregious example of a clip of catastrophic explosions being set to the song "I Just Wanna Fly" by Sugar Ray. Singer, "TED, "*see supra* note 36.

[145] Singer, WIRED, *Supra* note 10 at 85.

[146] *Robots at War: Drones and Democracy*, THE ECONOMIST (Oct. 1, 2010), http://www.economist.com/blogs/babbage/2010/10/robots_war.

6. INTERNATIONAL HUMANITARIAN SOMNAMBULISM?

Despite the length and complexity of this chapter, our conclusions are, in fact, quite modest. We have not speculated about the future uses of lethal autonomous robots. We have not argued in favor of or against their adoption. And we have not conjectured about whether their adoption might ever comport with the norms of international humanitarian law.

Instead, our analysis has focused solely on one means by which such technologies might be regulated. In considering the current military use of semi-autonomous machines and the case in favor of employing lethal autonomous robots in the future, we have examined international humanitarian law and have drawn some conclusions regarding its potential use in preventing unjustifiable military death and destruction by lethal autonomous robots.

We have suggested that it is crucial to recognize the philosophical underpinnings and implications of international humanitarian law's purportedly technology-neutral approach. Though this framework treats each technology under consideration as though it is neutral, the framework encourages and accommodates the development and use of emerging technologies. We believe that a failure to recognize and unpack the values embedded into the design of the framework itself, let alone those embedded in the robotic technologies under consideration, can lead to a mistaken and deterministic mode of thinking that fallaciously treats unjustifiable, lethal operations as though they are a military necessity. We have offered a possible explanation for how this might occur: when a disruptive technology changes the nature of what is possible, it also expands the scope of inclusion for what can prospectively be perceived of as "necessary" in carrying out military objectives.

Whether it turns out that we are right or wrong in offering this explanation, it is our hope that this chapter and its examination of how technologies imbue (and are imbued with) values creates space for alternative conceptions of regulating the military use of lethal autonomous robots.

Although we have neither discussed nor evaluated potential alternatives, many exist. Sector-specific frameworks are possible, as are conventions limiting or restricting the use of lethal autonomous weapons. There is precedent for international agreements to emerge from nongovernmental advocacy.[147] Some have advocated for a United Nations-sponsored

[147] The 1997 Ottawa Convention on the Prohibition of the Use, Stockpiling, Production, and Transfer of Anti-Personnel Mines and on Their Destruction was

international agreement.[148] Such suggestions are not implausible. In a 2010 report to the United Nations General Assembly, Christof Heyns, the UN Special Rapporteur on Extrajudicial Executions, questioned whether lethal force should ever be able to be truly automated.[149] Both Heyns and his immediate predecessor, Philip Alston, have recommended that the UN Secretary-General engage experts to evaluate possible parameters and conditions to restrain robot soldiers and recommend whether certain types of autonomous systems should be categorically unlawful.[150]

Currently agnostic with regard to which, if any, of these frameworks might emerge as most suitable, we have merely sought to demonstrate in this chapter that an exclusive reliance on the norms of international humanitarian law to regulate robotic warfare hazards increasing the trajectory of anti-humanitarian outcomes under the guise of military necessity. Permitting such a fictitious and fallacious treatment of lethal autonomous robots ultimately risks the very same somnambulism that would seek to promote humanitarian ends by taking humans out of the loop.

---

spurred by the dedicated efforts of civil society organizations worldwide. Notable activists within the movement included Diana, Princess of Wales, Canadian Foreign Minister Lloyd Axworthy, the International Campaign to Ban Landmines (ICBL, at www.icbl.org/intro.php), a global network that was awarded the Nobel Prize for its efforts to bring about the treaty. *See generally* Kenneth Anderson, *The Ottawa Convention Banning Landmines, the Role of International Non-governmental Organizations and the Idea of International Civil Society*, 11 *European Journal of International Law* 91, (2000).

148 Pagallo, *supra* note 39 at 322.

149 General Assembly GA/SHC/3986 Sixty-fifth General Assembly Third Committee 26th & 27th Meetings (AM & PM), http://www.un.org/News/Press/docs/2010/gashc3986.doc.htm.

150 Pagallo, *supra* note 39 at 321.