

2018

Autonomous Weapon Systems and the Limits of Analogy

Rebecca Crootof

University of Richmond - School of Law, rcrootof@richmond.edu

Follow this and additional works at: <https://scholarship.richmond.edu/law-faculty-publications>



Part of the [Military, War, and Peace Commons](#)

Recommended Citation

Rebecca Crootof, *Autonomous Weapon Systems and the Limits of Analogy*, 9 Harv. Nat'l Sec. J. 51 (2018).

This Article is brought to you for free and open access by the School of Law at UR Scholarship Repository. It has been accepted for inclusion in Law Faculty Publications by an authorized administrator of UR Scholarship Repository. For more information, please contact scholarshiprepository@richmond.edu.

ARTICLE

Autonomous Weapon Systems and the Limits of Analogy

Rebecca Crootof*

* Executive Director of the Information Society Project; Research Scholar and Lecturer in Law at Yale Law School. Many thanks to Charles Dunlap Jr. for an illuminating debate on the utility of new law in this context; to Claire Finkelstein, Duncan Macintosh, Jens David Ohlin, and William Ryan for clarifying questions and suggestions; and to Michael Adame, Douglas Bernstein, Christopher Izant, and Nathan Swire for extensive and thoughtful edits.

Copyright © 2018 by the President and Fellows of Harvard College and Rebecca Crootof.

Abstract

Autonomous weapon systems are often described either as more independent versions of weapons already in use or as humanoid robotic soldiers. In many ways, these analogies are useful. Analogies and allusions to popular culture make new technologies seem accessible, identify potential dangers, and buttress desired narratives. Most importantly from a legal perspective, analogical reasoning helps stretch existing law to cover developing technologies and minimize law-free zones.

But all potential analogies—weapon, combatant, child soldier, animal combatant—fail to address the legal issues raised by autonomous weapon systems, largely because they all misrepresent legally salient traits. Conceiving of autonomous weapon systems as weapons minimizes their capacity for independent and self-determined action, while the combatant, child soldier, and animal combatant comparisons overemphasize it. Furthermore, these discrete and embodied analogies limit our ability to think imaginatively about this new technology and anticipate how it might develop, thereby impeding our ability to properly regulate it.

We cannot simply graft legal regimes crafted to regulate other entities onto autonomous weapon systems. Instead, as is often the case when analogical reasoning cannot justifiably stretch extant law to answer novel legal questions, new supplemental law is needed. The sooner we escape the confines of these insufficient analogies, the sooner we can create appropriate and effective regulations for autonomous weapon systems.

Table of Contents

Introduction	54
I. Autonomous Weapon Systems.....	59
II. Common Analogies	62
A. Weapons	62
1. The law of weapons.....	62
2. The limits of the weapons analogy.....	64
B. Combatants	68
1. The law of combatants	68
2. The limits of the combatant analogy	70
III. Uncommon Analogies.....	73
A. Child Soldiers.....	73
1. The law of child soldiers	73
2. The limits of the child soldier analogy	75
B. Animal Combatants.....	76
1. The law of animal combatants	76
2. The limits of the animal combatant analogy	78
IV. The Limits of Analogy	79
A. Analogies Are Misleading	79
B. Analogies Are Constraining.....	80
Conclusion.....	82

Henry V: "An otter, Sir John. Why an otter?"
*Falstaff: "Why, she's neither fish nor flesh: a man knows not where to have her."*¹

Introduction

Autonomous weapon systems are usually depicted in one of two ways.² At times, they are discussed as more independent versions of weapons already in use—say, autonomous drones or smarter sea mines. Alternatively, they are portrayed as robotic soldiers or humanoid “killer robots,” conjuring images of the Terminator.³ These comparisons are useful: analogies to existing weaponry and allusions to popular culture make the incomprehensible seem accessible, highlight potential dangers, and buttress desired narratives.⁴ Analogical reasoning also helps stretch existing law to cover developing technologies and minimize law-free zones.⁵

¹ WILLIAM SHAKESPEARE, HENRY IV, PART I, act. 3, sc. 3.

² An “autonomous weapon system” is “a weapon system that, based on conclusions derived from gathered information and preprogrammed constraints, is capable of independently selecting and engaging targets.” Rebecca Crootof, *The Killer Robots Are Here: Legal and Policy Implications*, 36 CARDOZO L. REV. 1837, 1842 (2015); see also U.S. DEP’T OF DEF., DIRECTIVE 3000.09, AUTONOMY IN WEAPON SYSTEMS 13–14 (2012) (defining “autonomous weapon systems” as ones which, “once activated, can select and engage targets without further intervention by a human operator”). The term “autonomous weapon system” encourages the use of the weapon analogy, but I use it for the sake of consistency with the existing literature. A better term might simply be “autonomous agent.” Cf. Alessandro Guarino, *Autonomous Intelligent Agents in Cyber Offence*, in 5TH INTERNATIONAL CONFERENCE ON CYBER CONFLICT: PROCEEDINGS (K. Podins, J. Stinissen & M. Maybaum eds., 2013).

³ HUM. RTS. WATCH & INT’L HUM. RTS. CLINIC, HARV. L. SCH., LOSING HUMANITY: THE CASE AGAINST KILLER ROBOTS 38 (2012) [hereinafter LOSING HUMANITY] (suggesting that “emotionless robotic warriors” would be more likely to harm civilians and obey repressive autocrats); see also THE TERMINATOR (Orion Pictures 1984) (portraying a ruthless cyborg assassin, sent on a mission by Skynet, a self-aware artificial superintelligence intent on eliminating humanity).

⁴ An ongoing difficulty in discussions of how best to regulate autonomous weapon systems is that the term means different things to different people. See, e.g., Paul Scharre (@paul_scharre), TWITTER (Aug. 31, 2017, 10:00 AM), https://twitter.com/paul_scharre/status/903301434468442114 [<https://perma.cc/QH2M-JJMC>] (“[T]he term ‘autonomous weapon’ conjures up wildly different images for people. Some envision a Roomba w/ a gun. Others see Terminators.”). At present, there is still no agreed-upon international definition. UNIDIR, THE WEAPONIZATION OF INCREASINGLY AUTONOMOUS TECHNOLOGIES: CONCERNS, CHARACTERISTICS AND DEFINITIONAL APPROACHES 19–21 (2017) (outlining the three main definitional approaches). This lack of clarity encourages reliance on sometimes fanciful and inappropriate analogies.

⁵ Neil M. Richards & William D. Smart, *How Should the Law Think About Robots?*, in ROBOT LAW 3, 17 (Ryan Calo, A. Michael Froomkin & Ian Kerr eds., 2016) (“Particularly in the context

However, the weapon and combatant analogies for autonomous weapon systems are at odds with each other, insofar as they implicate distinct regulatory regimes. Weapons are inherently lawful or unlawful;⁶ combatants may act lawfully or unlawfully.⁷ Accordingly, the law of weapons regulates their physical design and capabilities, while the law governing combatants attempts to direct or constrain their behavior through a combination of training and accountability mechanisms.

Given this distinction, selecting the weapon or combatant analogy will predetermine the answers to many troubling legal questions, as that decision entails selecting between different standards for lawful use. Consider the customary requirement that parties to a conflict must distinguish between lawful targets (combatants, civilians directly participating in hostilities, and military objectives) and unlawful targets (civilians, surrendering or wounded combatants, and civilian objects).⁸ Autonomous weapon systems may be used in a discriminate manner,⁹ but they are not capable of independently distinguishing between lawful and unlawful targets, as would be expected of a combatant.¹⁰ Under the weapons rubric, autonomous weapon systems are lawful because they may be lawfully used; under the law regulating combatants, they (currently) cannot be expected to act lawfully.¹¹ Using one analogy, deploying autonomous weapon systems is lawful; using the other, it is not.

of changing technologies, the law almost always considers new technology as merely a new form of something else.”).

⁶ Granted, weapons that may be lawfully used in one manner or battlefield environment may be prohibited in others.

⁷ For the purposes of this paper, “combatants” entails members of the armed forces of a party to a conflict, all of whom are required to act in accordance with international humanitarian law. *See Rule 3. Definition of Combatants*, INT’L COMM. RED CROSS, https://ihl-databases.icrc.org/customary-ihl/eng/docs/v1_rul_rule3 (last visited Mar. 20, 2018) [<https://perma.cc/R4X-9MRG>]. The question of whether an individual is a “lawful” or “unlawful” combatant is relevant only in the context of determining whether that individual qualifies for prisoner of war status upon capture.

⁸ This customary rule is codified in multiple treaties, but most notably in Article 48 of the First Additional Protocol. *See Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of International Armed Conflicts (Protocol I)*, art. 48, *adopted* June 8, 1977, 1125 U.N.T.S. 3 [hereinafter *First Additional Protocol*].

⁹ Crootof, *supra* note 2, at 1874–76 (arguing that autonomous weapon systems may be used in compliance with the distinction requirement); Michael N. Schmitt, *Autonomous Weapon Systems and International Humanitarian Law: A Reply to the Critics*, HARV. NAT’L SEC. J. FEATURES (Feb. 5, 2013), at 10–13 (same).

¹⁰ Crootof, *supra* note 2, at 1873–74 (“At present, most agree that autonomous weapon systems are incapable of distinguishing between combatants and civilians. Doing so requires a complicated assessment of various factors, and there are many gray zones that bewilder even well-trained human soldiers.”); *see also* U.S. DEP’T OF DEF., *LAW OF WAR MANUAL* §6.5.9.3 (2015) (stating that the principle of proportionality applies to persons, not weapons) [hereinafter *DOD MANUAL*].

¹¹ Whether autonomous weapon systems will ever be capable of distinguishing between lawful and unlawful targets on par with human beings is a matter of hypothetical debate. Robotocist

Both those arguing for and those skeptical of a ban on autonomous weapon systems often shift fluidly between the weapon and combatant analogies to advance their preferred narratives, further muddling an already confused conversation.¹² Ban advocates utilize the combatant analogy to argue that “killer robots” will be incapable of complying with the law of targeting, cannot be held accountable for serious violations of international humanitarian law, and will be less moral than human soldiers (because they cannot empathize or show mercy).¹³ Meanwhile, they simultaneously rely on the weapon analogy when arguing that autonomous weapon systems can be successfully banned, just as other weapons have been banned.¹⁴ Ban skeptics analogize autonomous weapon systems to other weapons, arguing that they may be used in compliance with the law of armed conflict, possibly in more discriminate ways that better protect both combatants and civilians.¹⁵ But they also utilize the combatant analogy to argue that autonomous weapon systems may well be more humane than human soldiers, insofar as they will not make mistakes due to hunger, fatigue, illness, or boredom, and they will not act out of anger or a desire for revenge or renown.¹⁶

Granted, there are many circumstances where the best analogy for a given technology will shift depending on the legally salient characteristic of the issue being evaluated.¹⁷ It may be appropriate, for example, to consider the increasingly independent autonomous trucks now driving between Texas and California as “vehicles” in the context of evaluating whether speed limit laws apply, but as “drivers” or “employees” in evaluating which entity is liable should one hit a

Ronald Arkin suggests that robots will be able to comply with the distinction requirement in as few as ten years. Don Troop, *Robots at War: Scholars Debate the Ethical Issues*, CHRON. HIGHER EDUC. (Sept. 10, 2012), <http://chronicle.com/article/Moral-Robots-the-Future-of/134240>. Robotist Noel Sharkey does not believe that they will ever be able to do so. Noel E. Sharkey, *The Evitability of Autonomous Robot Warfare*, 94 INT’L REV. RED CROSS 787, 788–89 (2012).

¹² For example, “[m]any questions regarding accountability stem from a blurring of the approach to [autonomous weapon systems], an approach that moves from treating them as tools to treating them as moral or legal agents.” HEATHER ROFF & RICHARD MOYES, MEANINGFUL HUMAN CONTROL, ARTIFICIAL INTELLIGENCE, AND AUTONOMOUS WEAPONS 6 (2016), <http://www.article36.org/wp-content/uploads/2016/04/MHC-AI-and-AWS-FINAL.pdf> [<https://perma.cc/PP5P-2W7H>] (briefing paper prepared for the Informal Meeting of Experts on Lethal Autonomous Weapons Systems, UN Convention on Certain Conventional Weapons). In retrospect, I am also guilty of switching between analogies in my own writing. See Crootof, *supra* note 2, at 1866–68 (comparing autonomous weapon systems to both weapons and to human soldiers).

¹³ See, e.g., LOSING HUMANITY, *supra* note 3.

¹⁴ See, e.g., HUM. RTS. WATCH & INT’L HUM. RTS. CLINIC, HARV. L. SCH., PRECEDENT FOR PREEMPTION: THE BAN ON BLINDING LASERS AS A MODEL FOR A KILLER ROBOTS PROHIBITION (2015) [hereinafter PRECEDENT FOR PREEMPTION].

¹⁵ See, e.g., Schmitt, *supra* note 9, at 25.

¹⁶ See, e.g., Crootof, *supra* note 2, at 1867–68.

¹⁷ Cf. Jack Balkin, *The Path of Robotics Law*, 6 CALIF. L. REV. CIR. 45, 46 (2015) (“When we consider how a new technology affects law, our focus should not be on what is *essential* about the technology but on what features of social life the technology makes newly *salient*.”).

car.¹⁸ Similarly, it will often be possible to use analogical reasoning to stretch principles and rules developed in the weapons or combatant context to govern the use of autonomous weapon systems, with the selection between the two grounded in a reasoned analysis of why one analogy better addresses the issues at stake.¹⁹

But in some situations, neither the weapon nor the combatant analogy will provide a sensible or satisfying answer to a legal question, largely because both misrepresent what is fundamentally new about autonomous weapon systems. They are weapons, but weapons that may take independent actions.²⁰ They are independent actors, but actors that may be completely controlled in real time, either by their deployers or by hackers, simply by switching them from an “autonomous” to a “semi-autonomous” mode.

There are other entities that participate in armed conflicts—namely, children and animals—that are neither conventional weapons nor combatants. Like autonomous weapon systems, these nontraditional warfighters are capable of autonomous action and, by extension, they may sometimes act in unpredictable ways. But when their unexpected actions cause what appears to be a serious violation of international humanitarian law, they cannot be held individually liable under existing international criminal law, as they do not have the requisite *mens rea* to commit a war crime.²¹ Given the similarities between autonomous weapon systems, child soldiers, and animal combatants, it is tempting to think that we can look to those legal regimes for regulatory guidance.

Unfortunately, the solution is not that simple. The law of armed conflict has little to say about either child or animal combatants, rendering these analogies unhelpful in regulating autonomous weapon systems. Child soldiers are banned to protect children from the horrors of war; the same reasoning hardly applies to autonomous weapon systems. Meanwhile, the law of animal combatants is nearly nonexistent.

¹⁸ Alex Davies, *Self-Driving Trucks Are Now Delivering Refrigerators*, WIRED (Nov. 13, 2017), <https://www.wired.com/story/embark-self-driving-truck-deliveries/> [https://perma.cc/T8S9-XZC4].

¹⁹ Genevieve Lakier, *The Problem Isn't the Use of Analogies But the Analogies Courts Use*, KNIGHT FIRST AMENDMENT INST., https://knightcolumbia.org/content/problem-isnt-use-analogies-analogies-courts-use#/_ftnref14 (last visited May 4, 2018) [https://perma.cc/Q6ET-VTUN] (“[A]nalogies will prove useful only to the extent they are used thoughtfully, to illuminate the similarities and dissimilarities that matter for the purposes of the law.”).

²⁰ See, e.g. Alan L. Schuller, *At the Crossroads of Control: The Intersection of Artificial Intelligence in Autonomous Weapon Systems with International Humanitarian Law*, 8 HARV. NAT'L SEC. J. 379, 380 (2017) (“Advances in AI will likely produce [autonomous weapon systems] that are different *in kind* from existing weapon systems and thus require a fresh approach to evaluating [compliance with international humanitarian law].”).

²¹ See *infra* notes 100–101, 113, 140 and accompanying text.

Additionally, all of these potential analogies—weapon, combatant, child soldier, animal—are inappropriate and constraining. They either fail to capture a legally salient characteristic, or they imply the existence of a trait that is not actually there. Furthermore, by limiting our understanding of this new technology, these analogies impede our ability to appropriately regulate it.²² Rather than being a single, embodied unit, autonomous weapon systems will likely take a variety of forms, ranging from disembodied malware to networked systems of sensors and robots. These different forms and capabilities will affect how autonomous weapon systems can or should be regulated.

Many autonomous weapon systems in use today are relatively uncontroversial because they can be analogized to other weapons and regulated accordingly.²³ But the strain is already showing. It is not clear, for example, how existing law applies to weapon systems with in-field learning capabilities²⁴ or autonomous cyberoperations.²⁵ As increasingly autonomous weapon systems are developed and deployed, there will be additional situations where no analogy will be appropriate. As is often the case when analogical reasoning cannot justifiably

²² Cf. Duncan Hollis, *Re-Thinking the Boundaries of Law in Cyberspace: A Duty to Hack?*, in *CYBERWAR: LAW AND ETHICS FOR VIRTUAL CONFLICTS* (J. Ohlin et al. eds., 2015) (discussing issues with applying territorial boundary-based rules in cyber “space”).

²³ Crotoft, *supra* note 2, at 1873. Of course, all weapon-like autonomous weapon systems must be analogized to the appropriate weapon. An autonomous munition, an autonomous weapons platform, and an autonomous operational system will have fundamentally different capabilities, notwithstanding the fact that all are capable of independently selecting and engaging targets. See Michael Horowitz, *Why Words Matter: The Real World Consequences of Defining Autonomous Weapon Systems*, 30 *TEMPLE INT’L & COMP. L.J.* 85, 86 (2016) (proposing that new regulations distinguish between munition, platform, and operational autonomous weapon systems). Stationary and mobile weapon systems, embodied and software weapon systems, and short- and long-range weapon systems will have different reaches and ranges for the use of force—and different damage potentials should they act unpredictably. See PAUL SCHARRE, *AUTONOMOUS WEAPONS AND OPERATIONAL RISK* 12 (2016) https://s3.amazonaws.com/files.cnas.org/documents/CNAS_Autonomous-weapons-operational-risk.pdf?mtime=20160906080515 [<https://perma.cc/CJC9-6RX8>] (defining “damage potential” as “the amount of damage an autonomous system could do, if it failed to perform appropriately, before a human operator could take corrective action” and as depending “upon the inherent hazard of the system—the type of task being performed and the environment in which it is operating”).

²⁴ James Farrant & Christopher M. Ford, *Autonomous Weapons and Weapon Reviews: The UK Second International Weapon Review Forum*, 93 *INT’L L. STUD.* 389, 406–07 (2017).

²⁵ Experts disagree over whether Stuxnet—malware that destroyed 1,000 Iranian centrifuges used to enrich uranium—constitutes an “armed attack” justifying the use of responsive force. TALLINN MANUAL 2.0 ON THE INTERNATIONAL LAW APPLICABLE TO CYBER OPERATIONS ¶ 71 cmt. 10 (Michael N. Schmitt ed., 2017). Stuxnet had “a tremendous amount of autonomy,” but that autonomy was bounded by various safety features. PAUL SCHARRE, *ARMY OF NONE: AUTONOMOUS WEAPONS AND THE FUTURE OF WAR* 214–16 (2018); see also KERSTEIN VIGNARD, *THE WEAPONIZATION OF INCREASINGLY AUTONOMOUS TECHNOLOGIES: AUTONOMOUS WEAPON SYSTEMS AND CYBER OPERATIONS* 19 (2017) (arguing that those discussing the regulation of autonomous weapon systems should consider the relevance of increasingly autonomous cyber technologies).

stretch extant law to address novel legal questions raised by a new technology, new law is needed.

I. Autonomous Weapon Systems

Weapon systems can be roughly grouped into four categories of autonomy: inert, automated, semi-autonomous, and autonomous. Inert weapons are “objects requiring contemporaneous operation by a human being to be lethal,” and include everything from stones to advanced handheld firearms.²⁶ Automated weapon systems may be deployed long before they engage an unknown and unpredicted target, but they act in a predictable manner. A dumb landmine or tripwire sentry gun acts independently, but it does not “select” a target; rather, it responds predictably to a preset trigger.²⁷ A semi-autonomous weapon system, in contrast, has autonomous capabilities in functions relevant to target selection and engagement—but it cannot both select and engage targets independently.²⁸ Most “fire and forget” missiles and unmanned drones in operation today are semi-autonomous, as a human operator must take some affirmative action for a target to be selected or for a selected target to be engaged. Finally, an autonomous weapon system is capable of independently selecting and engaging targets based on conclusions derived from gathered information and preprogramed constraints.²⁹

Over thirty states already have “air, rocket, and missile defense systems with human-supervised autonomous modes,”³⁰ and states continue to develop increasingly autonomous weapon systems.³¹ The U.S. Aegis control system, operated in conjunction with the U.S. Phalanx Close In Weapons System (CIWS), is an autonomous weapon system, insofar as it has an operational mode that presumes human operators are incapacitated and allows it to independently identify and engage incoming anti-ship missiles and aircraft.³² The Israeli Harpy

²⁶ Crotoof, *supra* note 2, at 1864.

²⁷ *Id.* at 1857–63 (discussing what target selection and engagement entails, including the distinctions between specific and general targets and the relevance of sufficient time for affirmative human action); *id.* at 1864–65 (discussing automated weapons).

²⁸ *Id.* at 1865.

²⁹ *Id.* at 1855–56 (noting that such constraints may include general requirements grounded in the law of armed conflict to specific parameters of a given mission); SCHARRE, *supra* note 25, at 50 (“It is freedom, not intelligence, that defines an autonomous weapon system.”).

³⁰ SCHARRE, *supra* note 23, at 47.

³¹ *Cf.* Farrant & Ford, *supra* note 24, at 395–99 (discussing current, near-term, and long-term trends in autonomous weapon systems research and development).

³² John Pike, *MK 15 Phalanx Close-In Weapons System (CIWS)*, FED’N AM. SCIENTISTS, <http://www.fas.org/man/dod-101/syst/weaps/mk-15.htm> (last visited May 26, 2018) [<https://perma.cc/6SN8-FFP5>]; see also Gary E. Marchant et al., *International Governance of Autonomous Military Robots*, 12 COLUM. SCI. & TECH. L. REV. 272, 287 (2011) (describing the “casualty” setting, which allows the system to do “what it thinks is necessary to save the ship”).

Loitering Weapon, an airborne weapon that identifies and destroys enemy radar emitters, is another example of a fielded autonomous weapon system.³³

Autonomous weapon systems in use today act in largely predictable ways, either because they are operated in semi-autonomous modes or in relatively stable environments. For example, the South Korean SGR-A1 is a stationary, armed robot that allegedly has a fully autonomous setting, but South Korea maintains that it is used only in conjunction with a human operator and exclusively to monitor the relatively static Korean demilitarized zone.³⁴

It is impossible, however, to guarantee that autonomous weapon systems will always perform as expected. In addition to the more mundane sources of unpredictability extant in any weapon system, such as manufacturing defects and error rates, autonomous weapon systems are subject to additional kinds of malfunction.³⁵ They are necessarily complex systems, and the more complex their programming, the more opportunity there is for unforeseen interactions that cause unexpected results.³⁶ Elements in a system can interact in unanticipated ways that create or compound the harms of any single discrete failure.³⁷ Human-to-system interfaces introduce additional opportunities for confusion and catastrophic errors,³⁸ and it is nearly impossible to predict how one state's autonomous weapon systems will interact with allies' and enemies' systems.³⁹ Nor will it

³³ *Harpy Loitering Weapon*, ISR. AEROSPACE INDUS., http://www.iai.co.il/2013/36694-16153-en/Business_Areas_Land.aspx (last visited May 4, 2018).

³⁴ SCHARRE, *supra* note 25, at 104–05.

³⁵ In addition to malfunctions described in the text, autonomous weapon systems are also at risk of hacking and other malicious action. *See, e.g.*, MILES BRUNDIGE ET AL., *THE MALICIOUS USE OF ARTIFICIAL INTELLIGENCE: FORECASTING, PREVENTION, AND MITIGATION* 17–18 (2018) (“Today’s AI systems suffer from a number of novel unresolved vulnerabilities. . . . These vulnerabilities are distinct from traditional software vulnerabilities (e.g. bugger overflows) and demonstrate that while AI systems can exceed human performance in many ways, they can also fail in ways that a human never would.”).

³⁶ SCHARRE, *supra* note 23, at 13; *see also* SCHARRE, *supra* note 25, at 210 (quoting Bradford Tousley as observing that automated stock trading—with its attendant unavoidable “flash crashes”—is a “great analogy” for the challenges of automation in military applications”).

³⁷ Stephanie Carvin, *Normal Autonomous Accidents: What Happens When Killer Robots Fail?* 1 (Mar. 1, 2017) (unpublished manuscript), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3161446 (discussing the relevance of “Normal Accident Theory (and its critics) to explore the concerns raised over [lethal autonomous weapon systems] through the lens of system failure”); SCHARRE, *supra* note 23, at 25 (discussing “normal accidents” theory).

³⁸ M.C. Elish, *Moral Crumple Zones: Cautionary Tales in Human-Robot Interaction* (Mar. 20, 2016) (unpublished manuscript), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2757236 (discussing situations where miscommunications between human beings and machine systems resulted in preventable accidents).

³⁹ Carvin, *supra* note 37, at 16 (“[W]e should very well expect serious challenges when [lethal autonomous weapon systems] begin to regularly encounter one another.”); Heather Roff, *The Strategic Robot Problem: Lethal Autonomous Weapons in War*, 13 *J. MIL. ETHICS* 211, 218 (2014)

always be possible to determine when a system is malfunctioning, as “systems are not always good at communicating what is going on when something goes wrong.”⁴⁰ Finally, to the extent autonomous weapon systems rely on artificial neural networks—digital networks that “learn” from data, including past actions, to develop their own conclusions and rules⁴¹—the reasons for a given action may be as inscrutable as the action was unexpected.⁴²

In short, autonomous weapon systems have the capacity for independent and thus inherently unpredictable action, which both distinguishes them from other types of weapons and raises a host of legal questions.⁴³ What kind of legal review is sufficient for an entity with in-field learning capabilities?⁴⁴ What constitutes “meaningful human control” over an autonomous weapon system?⁴⁵ When and how should someone be held accountable if an autonomous weapon system’s actions result in a serious violation of international humanitarian law?⁴⁶

(noting that “international or allied joint forces activities are [already] plagued by operational difficulties due to the joining of different militaries”).

⁴⁰ Carvin, *supra* note 37, at 9 (discussing how malfunctioning systems can provide misleading information and multifaceted systems make it difficult to identify problems).

⁴¹ DAVID KRIESEL, A BRIEF INTRODUCTION TO NEURAL NETWORKS 7 (2007), http://www.dkriesel.com/_media/science/neuronalenetze-en-zeta2-2col-dkrieselcom.pdf [<https://perma.cc/9PWA-QFJN>].

⁴² SCHARRE, *supra* note 23, at 15; *see also* Joel Lehman et al., The Surprising Creativity of Digital Evolution: A Collection of Anecdotes from the Evolutionary Computation and Artificial Life Research Communities (Mar. 29, 2018) (unpublished manuscript), <https://arxiv.org/pdf/1803.03453.pdf> [<https://perma.cc/7A5H-64HM>] (collecting stories of evolving algorithms acting unexpectedly). AlphaGo, for example, is an algorithm that mastered the board game Go by repeatedly playing itself. Expert Go players have described its strategies as brilliant and incomprehensible. *See* Dawn Chan, *The AI That Has Nothing to Learn From Humans*, ATLANTIC (Oct. 20, 2017), https://www.theatlantic.com/technology/archive/2017/10/alphago-zero-the-ai-that-taught-itself-go/543450/?utm_source=atfb [<https://perma.cc/DJE6-WQ85>] (quoting experts as saying the moves were “how I imagine games from far in the future,” “Go from an alternate dimension,” and “alien”).

⁴³ Of course, autonomous weapon systems also raise a host of moral, strategic, and political questions as well. This paper, however, is primarily concerned with how analogies are employed to (sometimes inappropriately) stretch the law to cover this new technology.

⁴⁴ *See* Farrant & Ford, *supra* note 24, at 406–07 (discussing Article 36 legal reviews for continuously learning systems).

⁴⁵ *See, e.g.* Michael C. Horowitz & Paul Scharre, *Meaningful Human Control in Weapon Systems: A Primer* 7 (Ctr. for New Am. Sec., Working Paper No. 031315, 2015), https://www.files.ethz.ch/isn/189786/Ethical_Autonomy_Working_Paper_031315.pdf [<https://perma.cc/KY8C-FHX2>] (describing two main schools of thought regarding the principle’s legal status); *see also* Rebecca Crootof, *A Meaningful Floor for ‘Meaningful Human Control’*, 30 TEMPLE INT’L & COMP. L.J. 53, 58–60 (2016) (noting benefits of retaining imprecision in international standards governing new technology but arguing for an interpretative floor for the concept of “meaningful human control” grounded on existing humanitarian protections).

⁴⁶ *See, e.g.*, Rebecca Crootof, *War Torts: Accountability for Autonomous Weapons*, 164 U. PA. L. REV. 1347, 1376–77 (2016) (discussing when individuals can and cannot be held directly

In the attempt to answer these questions, scholars and advocates are relying on familiar but inadequate—and therefore misleading—analogies.

II. Common Analogies

Autonomous weapon systems are commonly analogized to weapons and combatants, implicating two distinct legal regimes with different foundational assumptions. This Section reviews basic tenants of both legal regimes and identifies legal questions relevant to the governance of autonomous weapon systems that neither regime can adequately address absent the creation of new, supplementary rules.

A. Weapons

1. The law of weapons

It is widely recognized that “[i]n any armed conflict, the right of the Parties to the conflict to choose methods or means of warfare is not unlimited.”⁴⁷ Both customary international law and various treaties circumscribe which weapons may be lawfully fielded.⁴⁸

Weapons that cause unnecessary injury or suffering are prohibited. This tech-neutral prohibition was first articulated by the European powers in the Saint Petersburg Declaration of 1868, which forbade the use of weapons “which uselessly aggravate the sufferings of disabled men, or render their death inevitable” as “contrary to the laws of humanity.”⁴⁹ It has since been reiterated in numerous treaties and is now recognized as customary international law.⁵⁰

The customary prohibition on indiscriminate attacks implies that inherently indiscriminate weapons—those that cannot be directed at a military objective or whose effects cannot be controlled—are also per se unlawful.⁵¹ As is

criminally liable for the acts of an autonomous weapon system); *id.* at 1386 (arguing that states should be held strictly liable for the harms caused by their autonomous weapon systems).

⁴⁷ Convention Respecting the Laws and Customs of War on Land art. 22, Oct. 18, 1907, 36 Stat. 2277, 1 Bevans 631; *see also* First Additional Protocol, *supra* note 8, art. 35(1).

⁴⁸ Customary international law derives from “evidence of a general practice accepted as law.” Statute of the International Court of Justice art. 38, June 26, 1945, T.S. No. 993. While treaty law is binding only on states party to the treaty, customary international law binds all states (subject to a limited “opt-out” exception).

⁴⁹ Declaration Renouncing the Use, in Time of War, of Certain Explosive Projectiles, Nov. 29–Dec. 11, 1868, 18 Martens Nouveau Recueil (ser. 1) 474, 138 Consol. T.S. 297.

⁵⁰ *Rule 70. Weapons of a Nature to Cause Superfluous Injury or Unnecessary Suffering*, INT’L COMM. OF THE RED CROSS, http://www.icrc.org/customary-ihl/eng/docs/v1_rul_rule70 (last visited May 4, 2018) [<https://perma.cc/KYM8-NVMB>].

⁵¹ *Rule 71. Weapons That Are by Nature Indiscriminate*, INT’L COMM. OF THE RED CROSS, https://www.icrc.org/customary-ihl/eng/docs/v1_rul_rule71 (last visited May 4, 2018) [<https://perma.cc/G8V3-W942>].

often the case with tech-neutral rules, what constitutes an inherently indiscriminate weapon has evolved with new technological developments. For example, “many of the gravity bombs designed for release from high altitudes that were dropped during World War II would today be characterized as indiscriminate.”⁵²

While less established, weapons that “are intended, or may be expected, to cause widespread, long-term and severe damage to the natural environment”⁵³ are likely also unlawful. This rule was codified in the 1977 First Additional Protocol to the Geneva Conventions.⁵⁴ Since then, the prohibition on the *intentional* destruction of the natural environment has been recognized by many states as customary international law,⁵⁵ which implies that weapons intended to cause such destruction are prohibited.

Finally, there are a number of tech-specific weapons bans and regulations, usually enacted through treaties, which have enjoyed varying degrees of success.⁵⁶ The ban on “laser weapons specifically designed, as their sole combat function or as one of their combat functions, to cause permanent blindness to unenhanced vision”⁵⁷ may well be the most successful ban of all time, as there is little controversy over what it prohibits and there are no recorded violations.⁵⁸ Other treaty bans, like the prohibition on the use of chemical weapons, are now

⁵² Schmitt, *supra* note 9, at 10 n.29.

⁵³ *Rule 45. Causing Serious Damage to the Natural Environment*, INT’L COMM. OF THE RED CROSS, https://www.icrc.org/customary-ihl/eng/docs/v1_rul_rule45 (last visited May 4, 2018) [<https://perma.cc/5FG3-KHAJ>] [hereinafter *ICRC Rule 45*].

⁵⁴ First Additional Protocol, *supra* note 8, art. 35(3).

⁵⁵ *ICRC Rule 45*, *supra* note 53 (discussing state practice supporting and limiting this claim).

⁵⁶ For a list of oft-cited bans of specific weapons in the autonomous weapon systems debate and common wisdom regarding their respective successes and failures, see Crootof, *supra* note 2, appendix (discussing various attempted and enacted weapons bans, including crossbows, aerial bombardment, submarines, nuclear weapons, cluster munitions, anti-personnel landmines, biological weapons, chemical weapons, and permanently blinding lasers).

⁵⁷ Protocol on Blinding Laser Weapons (Protocol IV) art. 1, Oct. 13, 1995, S. Treaty Doc. No. 105-1, 2024 U.N.T.S. 167.

⁵⁸ Crootof, *supra* note 2, at 1915. Given its success, advocates for a ban on autonomous weapon systems often tout the ban on blinding laser weapons to argue that preemptive weapons bans can be highly effective. See generally, e.g., PRECEDENT FOR PREEMPTION, *supra* note 14; MINES ACTION CANADA, LESSONS FROM PROTOCOL IV ON BLINDING LASER WEAPONS FOR THE CURRENT DISCUSSIONS ABOUT AUTONOMOUS WEAPONS (2014), <https://bankillerrobotscanada.files.wordpress.com/2014/05/international-piv-memo-final.pdf> [<https://perma.cc/FA8Z-F5GF>]. However, permanently blinding laser weapons share few characteristics with autonomous weapon systems, rendering the comparison inapt. See Rebecca Crootof, *Why the Ban on Permanently Blinding Lasers is Poor Precedent for a Prohibition on Autonomous Weapon Systems*, LAWFARE (Nov. 24, 2015), <https://www.lawfareblog.com/why-prohibition-permanently-blinding-lasers-poor-precedent-ban-autonomous-weapon-systems> [<https://perma.cc/YLU2-GNYH>].

recognized as customary international law.⁵⁹ However, other attempts to prohibit the use of certain weapons were less successful: the ban on the use of crossbows failed immediately and dramatically,⁶⁰ while the ban on aerial bombardment did not outlast the invention of the airplane.⁶¹

To ensure that fielded weapons may be lawfully used, states are charged with conducting legal reviews of new weapon designs.⁶² Of course, any lawful weapon may be used in an unlawful manner, which is why international humanitarian law also regulates the behavior of individuals in armed conflict.

2. The limits of the weapons analogy

Analogizing autonomous weapon systems to conventional weapons simplifies some legal analyses, but that simplicity can prove misleading. Certainly, an autonomous weapon system is not inherently unlawful if it (1) will not cause superfluous injury or unnecessary suffering, (2) is capable of being used in a discriminate manner, (3) is not intended to or is not expected to cause widespread and severe environmental damage, and (4) does not employ weapons that are specifically prohibited (like poison or permanently blinding lasers). Autonomous weapon systems currently in use have cleared this low bar—in part because, as noted earlier, they are being used in predictable environments or in semi-autonomous modes.

However, the weapon analogy complicates other legal analyses. First, this analogy raises the tricky question of what constitutes a sufficient legal review for increasingly autonomous weapon systems.⁶³ Given that its actions will be determined in part by stimuli from its environment, an autonomous weapon system will need to be tested in a variety of different virtual and real scenarios and only cleared for use in sufficiently similar real world circumstances.⁶⁴ For

⁵⁹ See, e.g., Rule 74. *Chemical Weapons*, INT'L COMM. OF THE RED CROSS, http://www.icrc.org/customary-ihl/eng/docs/v1_cha_chapter24_rule74 (last visited May 4, 2018) [<https://perma.cc/7J45-URPC>].

⁶⁰ W.T. Mallison, Jr., *The Laws of War and the Juridical Control of Weapons of Mass Destruction in General and Limited Wars*, 36 GEO. WASH. L. REV. 308, 316 (1967).

⁶¹ Arthur K. Kuhn, *The Beginnings of an Aërial Law*, 4 AM. J. INT'L L. 109, 119–20 (1910) (suggesting that states' failure to renew the ban was likely linked to their interest in exploring the military applications of airplanes).

⁶² This obligation, codified in the First Additional Protocol, *supra* note 8, art. 36, is also binding on all states under customary international law, see Kathleen Lewand et al., *A Guide to the Legal Review of New Weapons, Means and Methods of Warfare: Measures to Implement Article 36 of Additional Protocol I of 1977*, 88 INT'L REV. RED CROSS 931, 933 (2006) (noting that the customary and treaty prohibitions on certain kinds of weapons imply that states are obliged to conduct some form of review to avoid fielding unlawful weapons).

⁶³ Cf. Farrant & Ford, *supra* note 24, at 395–99.

⁶⁴ Cf. U.S. DEP'T OF DEF., DEF. SCI. BD., SUMMER STUDY ON AUTONOMY 29 (2016) (recommending “a combination of modeling and simulation to explore thousands of text cases,

some systems—say, ones designed only to be used in outer space, where there is limited variation—that will be relatively easy. For others, like mobile land systems that might be deployed in dramatically different environments, it will be difficult to ensure that a system’s training is adequate. Accordingly, there is a risk that, absent international rules specifically governing autonomous weapon systems, states’ idiosyncratic interpretations of tech-neutral principles will result in the fielding of unlawful weapons.⁶⁵

Second, there is no precedent for evaluating weapon systems with the capacity for in-field learning.⁶⁶ Traditionally, legal reviews presume that any significant change in software or hardware will be identified and evaluated before the altered weapon is fielded. But weapon systems with certain kinds of artificial intelligence might be able to evolve in response to environmental stimuli, such that a fielded autonomous weapon system will eventually be significantly different from the system originally approved. This problem might be addressed in various ways: by short-term deployments, by regularly scheduled evaluations and reboots, or by completely prohibiting in-field learning (at least with regard to target selection and engagement tasks). All of these solutions, however, would likely require new domestic guidelines, if not international standards.

Granted, the legal review requirement already suffers from significant noncompliance. Most states do not publicly acknowledge that they conduct legal reviews of new weaponry,⁶⁷ and legal reviews have been critiqued for being perfunctory or for being easily evaded by presenting a new weapon as merely a minor modification of an approved precursor.⁶⁸ A silver lining is that the international conversation about how best to regulate autonomous weapon systems has the potential to highlight these issues and contribute to the development of widely-accepted standards for conducting legal reviews. Already,

statistically measuring system performance against the desired standard, then doing real world testing of the system to ensure that the modeled and real-world behavior match for corner cases that span the range of system performance”).

⁶⁵ ARTICLE 36, ARTICLE 36 REVIEWS AND ADDRESSING LETHAL AUTONOMOUS WEAPONS SYSTEMS 2–3 (2016) (briefing paper for delegates at the Convention on Certain Conventional Weapons Meeting of Experts on Lethal Autonomous Weapons Systems).

⁶⁶ Farrant & Ford, *supra* note 24, at 406–07; *see also* Ryan Calo, *Robotics and the Lessons of Cyberlaw*, 103 CALIF. L. REV. 513, 538–45 (2015) (discussing robotic “emergence,” the ability of robotic systems to adapt to circumstances and “learn” from mistakes).

⁶⁷ Farrant & Ford, *supra* note 24, at 401.

⁶⁸ For example, modifications to the US B61 nuclear bomb have sparked debate over whether it constitutes a new weapon. Julian Borger, *America’s New, More ‘Usable’, Nuclear Bomb in Europe*, GUARDIAN (Nov. 10, 2015), <https://www.theguardian.com/world/julian-borger-global-security-blog/2015/nov/10/americas-new-more-usable-nuclear-bomb-in-europe> [<https://perma.cc/KTX4-659S>].

a few states have been more transparent about their internal review procedures in discussions of regulating autonomous weapon systems.⁶⁹

Third, after an attack, it is unlikely that anyone will be held criminally liable or that victims will be compensated for the unpredictable but injurious actions of autonomous weapon systems. International criminal law requires that a person act “willfully”—which is usually understood as acting intentionally or recklessly—for criminal liability to attach.⁷⁰ Of course, someone who intentionally or recklessly uses an autonomous weapon systems to commit serious violations of international humanitarian law will be criminally liable,⁷¹ and states may voluntarily compensate the injured.⁷² But if an individual fields an autonomous weapon system that unexpectedly destroys a hospital, downs a passenger jet, or otherwise takes action that would appear to be a war crime, the individual deployer will not be criminally liable.⁷³ The destructive consequences of the malfunction will be considered an accident, a tragic but not unexpected side effect of war, and the harm will lie where it falls.⁷⁴

Fourth, and more generally, autonomous weapon systems allow for greater temporal and geographic distance between a human deployer’s decision to use lethal force and the consequence of that decision. Thus, analogizing autonomous

⁶⁹ ARTICLE 36, *supra* note 65, at 1, 5 n.2 (noting that Australia, Germany, the Netherlands, Russia, the United States, and the United Kingdom have provided outlines of their weapons review procedures in the context of discussing autonomous weapon systems).

⁷⁰ *See, e.g.*, Rome Statute of the International Criminal Court art. 30(1), July 17, 1998, 2187 U.N.T.S. 90 (“[A] person shall be criminally responsible and liable for punishment . . . only if the material elements are committed with intent and knowledge.”); *id.* art. 30(3) (defining “knowledge” as entailing “awareness that a circumstance exists or a consequence will occur in the ordinary course of events”); *see also* Prosecutor v. Blaškić, Case No. IT-95-14-T, Judgment, ¶ 152 (Int’l Crim. Trib. for the Former Yugoslavia Mar. 3, 2000) (“[T]he *mens rea* constituting all the [grave breaches of the Geneva Conventions] includes both guilty intent and recklessness which may be likened to serious criminal negligence.”).

⁷¹ Crotoft, *supra* note 46, at 1376–77 (describing these as the “easy cases” for evaluating accountability).

⁷² *See generally* John Fabian Witt, *Form and Substance in the Law of Counterinsurgency Damages*, 41 LOY. L.A. L. REV. 1455 (2008) (discussing the role of voluntary condolence payments in U.S. military strategy).

⁷³ Crotoft, *supra* note 46, at 1377 (“If an autonomous weapon system is merely another weapon in a state’s arsenal, its deployer will be liable only if she intended or foresaw the reasonable likelihood of civilian harm and nonetheless used the weapon system.”).

⁷⁴ Some have argued that this lack of accountability may have the perverse effect of encouraging commanders to inappropriately deploy autonomous weapon systems. HUM. RTS. WATCH & INT’L HUM. RTS. CLINIC AT HARV. L. SCH., MIND THE GAP: THE LACK OF ACCOUNTABILITY FOR KILLER ROBOTS 24 (2015) [hereinafter MIND THE GAP]; Peter Asaro, *On Banning Autonomous Weapon Systems: Human Rights, Automation, and the Dehumanization of Lethal Decision-Making*, 94 INT’L REV. RED CROSS 687, 701 (2012); Heather M. Roff, *Killing in War: Responsibility, Liability, and Lethal Autonomous Robots*, in ROUTLEDGE HANDBOOK OF ETHICS AND WAR 352, 357–58 (Fritz Allhoff, Nicholas G. Evans & Adam Henschke eds., 2013).

weapon systems to other weapons raises the question of what constitutes an appropriate level of human oversight or involvement in attacks generally.⁷⁵ To be sure, other existing technologies—like anti-personnel landmines, fire-and-forget missiles, and the drones used for targeted killings—already test the presumption that there is a direct link between a weapon’s deployer and its target, and these technologies are often critiqued (among other things) for permitting unacceptable detachment from the act of killing.⁷⁶ But autonomous weapon systems permit even greater degrees of temporal and geographic distance, making this latent issue even more salient.⁷⁷

While the existing law of weaponry will govern most uses of contemporary autonomous weapon systems, this new technology’s capacity for independent action raises issues that simply are not currently addressed by this legal regime. The weapons review process does not have widely accepted procedures for evaluating the lawfulness of independent or learning weapon systems, international criminal law does not hold individuals accountable for the unpredictable actions of autonomous weapon systems, and it is unclear whether there is customary international law requiring a certain level of human oversight or involvement in attacks. Given their ability to act as independent and unpredictable agents, it is tempting to analogize autonomous weapon systems to another autonomous warfighter—human combatants.

⁷⁵ ROFF & MOYES, *supra* note 12, at 6.

⁷⁶ See, e.g., Mary Ellen O’Connell, *Banning Autonomous Killing: The Legal and Ethical Requirement that Humans Make Near-Time Lethal Decisions*, in THE AMERICAN WAY OF BOMBING: CHANGING ETHICAL AND LEGAL NORMS, FROM FLYING FORTRESSES TO DRONES 224, 234 (Matthew Evangelista & Henry Shue eds., 2014) (arguing that a new norm of international law, requiring a close temporal space between force deployment and target engagement, is necessary to “keep a human conscience” in the decision to use lethal force). Of course, one of the primary aims of weapons development is to reduce risks to one’s forces by increasing geographic and temporal distances between them and dangerous targets. Jens David Ohlin, *Remoteness and Reciprocal Risk*, in RESEARCH HANDBOOK ON REMOTE WARFARE (Jens David Ohlin ed., 2017); Oren Gross, *The New Way of War: Is There a Duty to Use Drones?*, 67 FLA. L. REV. 1, 33 (2015). This tension—between the bent towards weapons that allow for distance and weapons that are not insufficiently precise—is a central one in international humanitarian law. Michael N. Schmitt, *Military Necessity and Humanity in International Humanitarian Law: Preserving the Delicate Balance*, 50 VA. J. INT’L L. 795 (2010).

⁷⁷ The resulting conversation over what constitutes “meaningful human control” in an attack is still in its infancy. Crootof, *supra* note 45, at 54 (observing that the broad support for the principle “comes at a familiar legislative cost: there is no consensus as to what ‘meaningful human control’ actually requires”). However, some argue that a formal norm is coalescing or already exists. Peter Asaro, *Jus Nascendi, Robotic Weapons and the Martens Clause*, in ROBOT LAW, *supra* note 5, 367, 368 (“[T]here is already an emerging norm concerning meaningful human control over the targeting of weapons and the use of violent force . . .”).

B. Combatants

1. The law of combatants

Generally, members of the armed forces, associated militias, and associated volunteer corps are considered combatants,⁷⁸ under certain (and debated) conditions, members of organized armed groups and civilians directly participating in hostilities may also be considered combatants.⁷⁹ In recognition of the inherent autonomy of combatants—the ability to observe, reflect, and choose a course of action⁸⁰—international humanitarian law attempts to regulate their behavior by, first, using training to help individuals internalize international humanitarian law and, second, through the deterrent power of punishment.

There are myriad treaties and customary international laws governing combatants' actions in armed conflicts,⁸¹ but the law of targeting—which might be characterized as the law of when, how, and against whom weapons may be employed in attacks—is of particular relevance to this discussion. The law of targeting is grounded on the principles of military necessity and humanity, which manifest in the application of the distinction, proportionality, and feasible precautions requirements.⁸² The distinction requirement obliges parties to a conflict to distinguish between lawful and unlawful targets.⁸³ The proportionality requirement prohibits any attack that “may be expected to cause incidental loss of civilian life, injury to civilians, damage to civilian objects, or a combination thereof, which would be excessive in relation to the concrete and direct military advantage anticipated.”⁸⁴ Lastly, the feasible precautions requirement mandates that a commander shall (1) “do everything feasible to verify that the objectives to

⁷⁸ See, e.g., Geneva Convention Relative to the Treatment of Prisoners of War art. 4(A), Aug. 12, 1949, 6 U.S.T. 3316, 75 U.N.T.S. 135 [hereinafter Third Geneva Convention]; First Additional Protocol, *supra* note 8, arts. 43–44.

⁷⁹ See generally INT'L COMM. OF THE RED CROSS, INTERPRETIVE GUIDANCE ON THE NOTION OF DIRECT PARTICIPATION IN HOSTILITIES UNDER INTERNATIONAL HUMANITARIAN LAW (2009). For the purposes of this paper, I do not distinguish between privileged or unprivileged combatants.

⁸⁰ “Autonomy,” as a philosophical concept, is far from well defined. See, e.g., John Christman, *Autonomy in Moral and Political Philosophy*, in STANFORD ENCYCLOPEDIA OF PHILOSOPHY (Edward N. Zalta ed., 2015), <http://plato.stanford.edu/entries/autonomy-moral> (discussing conceptual variations and controversies).

⁸¹ Including prohibitions on deception, rules regarding communication with adversarial forces, and requirements concerning the treatment of detainees.

⁸² Schmitt, *supra* note 76, at 803–05.

⁸³ First Additional Protocol, *supra* note 8, art. 48.

⁸⁴ *Id.*, art. 51(5)(b). Such determinations are evaluated under a “reasonable commander” standard. See Prosecutor v. Galić, Case No. IT-98-29-T, Judgment, ¶ 58 (Int'l Crim. Trib. for the Former Yugoslavia Dec. 5, 2003) (“In determining whether an attack was proportionate it is necessary to examine whether a reasonably well-informed person in the circumstances of the actual perpetrator, making reasonable use of the information available to him or her, could have expected excessive civilian casualties to result from the attack.”).

be attacked are neither civilians nor civilian objects and are not subject to special protection but are military objectives”; (2) “take all feasible precautions in the choice of means and methods of attack with a view to avoiding, and in any event to minimizing, incidental loss of civilian life, injury to civilians and damage to civilian objects”; and (3) select the objective that “may be expected to cause the least danger to civilian lives and to civilian objects” when choosing among objectives with similar military advantages.⁸⁵ As with the seemingly tech-neutral rules of weapon design, new technological developments have sometimes raised confusion in how these rules should be interpreted and applied.⁸⁶

To ensure that combatants internalize these and other foundational norms of international humanitarian law, the 1949 Geneva Conventions require state parties to train members of their armed forces in the law and to disseminate the Conventions “as widely as possible.”⁸⁷ The International Committee of the Red Cross (ICRC) has identified these requirements as customary international law.⁸⁸

⁸⁵ First Additional Protocol, *supra* note 8, art. 57; *see also* DOD MANUAL, *supra* note 10, § 6.5.9.3 (noting that, in the context of autonomous weapon systems, feasible precautions might “include monitoring the operation of the weapon system or programming or building mechanisms for the weapon to deactivate automatically after a certain period of time”).

⁸⁶ *Cf.* Crootof, *supra* note 46, at 1370–71 (“As weapon systems become more accurate and require less lethal force for effectiveness, what is considered a ‘proportionate’ level of collateral damage has narrowed dramatically.”); Jens David Ohlin, *Is Jus in Bello in Crisis?*, 11 J. INT’L CRIM. JUST. 27, 28 (2013) (“If the use of drone technology against non-state actors (NSAs) has made anything clear, it is this: there is far less agreement regarding the application of core principles of international humanitarian law . . . and international criminal law . . . than previously thought . . .”).

⁸⁷ Geneva Convention for the Amelioration of the Condition of the Wounded and Sick in Armed Forces in the Field art. 47, Aug. 12, 1949, 6 U.S.T. 3114, 75 U.N.T.S. 31 [hereinafter First Geneva Convention]; Geneva Convention for the Amelioration of the Condition of Wounded, Sick and Shipwrecked Members of Armed Forces at Sea art. 48, Aug. 12, 1949, 6 U.S.T. 3217, 75 U.N.T.S. 85 [hereinafter Second Geneva Convention]; Third Geneva Convention, *supra* note 78, art. 127; Geneva Convention Relative to the Protection of Civilian Persons in Time of War art. 144, Aug. 12, 1949, 6 U.S.T. 3516, 75 U.N.T.S. 287 [hereinafter Fourth Geneva Convention]. The obligation to disseminate is reiterated in the two 1977 Additional Protocols. First Additional Protocol, *supra* note 8, art. 83; Protocol Additional (II) to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of Non-International Armed Conflicts art. 19, June 8, 1977, 1125 U.N.T.S. 609.

⁸⁸ *Rule 142. Instruction in International Humanitarian Law within Armed Forces*, INT’L COMM. OF THE RED CROSS, https://www.icrc.org/customary-ihl/eng/docs/v1_rul_rule142 (last visited May 4, 2018) [<https://perma.cc/E4KF-WL65>] (“States and parties to the conflict must provide instruction in international humanitarian law to their armed forces.”); *Rule 143. Dissemination of International Humanitarian Law Among the Civilian Population*, INT’L COMM. OF THE RED CROSS, https://www.icrc.org/customary-ihl/eng/docs/v1_rul_rule143 (last visited May 4, 2018) [<https://perma.cc/JR9S-XX3A>].

However, in recognition that training will not completely prevent autonomous combatants from committing serious violations,⁸⁹ international humanitarian law also relies on the deterrent power of punishment. State parties to the Geneva Conventions are required to search for and try individuals that have allegedly committed grave breaches of the Conventions,⁹⁰ and the ICRC has identified the duty to prosecute and punish war crimes as customary international law.⁹¹ Nor are only those individuals who commit grave breaches potentially liable; it is generally understood that a superior may be liable under the customary doctrine of “command responsibility” if “she exercises effective control over a subordinate, knows of or has reason to know of the subordinate’s actual or intended criminal acts, and fails to take necessary and reasonable measures to prevent or punish them.”⁹² The “prevent or punish” requirement emphasizes the import of minimizing violations while implicitly acknowledging the impossibility of always being able to prevent autonomous agents from acting unlawfully.

2. The limits of the combatant analogy

Employing the combatant analogy sidesteps the problem of what constitutes adequate legal review, but it raises other issues regarding appropriate training and the utility of threatened punishment.

How are autonomous weapon systems to be adequately “trained” in the principles governing targeting and other law of armed conflict rules?⁹³ Trainings developed for human beings assume that the individual has common sense, the flexibility to apply general rules to new situations, and a basic morality that

⁸⁹ See, e.g., David Lloyd Roberts, *Training the Armed Forces to Respect International Humanitarian Law: The Perspective of the ICRC Delegate to the Armed and Security Forces of South Asia*, 319 INT’L REV. RED CROSS 433, 434 (1997) (citing Yugoslavia as an “example of a failure to implement the law [of armed conflict]—despite apparent good training”).

⁹⁰ First Geneva Convention, *supra* note 87, art. 49; Second Geneva Convention, *supra* note 87, art. 50; Third Geneva Convention, *supra* note 78, art. 129; Fourth Geneva Convention, *supra* note 87, art. 146.

⁹¹ *Rule 158. Prosecution of War Crimes*, INT’L COMM. OF THE RED CROSS, https://www.icrc.org/customary-ihl/eng/docs/v1_rul_rule158 (last visited May 4, 2018) [<https://perma.cc/23B3-6YFA>] (citing military manuals, national legislation, and official statements).

⁹² Crootof, *supra* note 46, at 1378 n.172 (citing Statute of the Special Court for Sierra Leone art. 6(3), Jan. 16, 2002, 2178 U.N.T.S. 145; Rome Statute, *supra* note 70, art. 28; Statute of the International Criminal Tribunal for Rwanda art. 6(3), Nov. 8, 1994, 33 I.L.M. 1598; Updated Statute of the International Criminal Tribunal for the Former Yugoslavia art. 7(3), May 25, 1993, 32 I.L.M. 1192; First Additional Protocol, *supra* note 8, arts. 86–87; Prosecutor v. Delalić, Case No. IT-96-21-T, Judgment, ¶ 346 (Int’l Crim. Trib. for the Former Yugoslavia Nov. 16, 1998); ANTONIO CASSESE, INTERNATIONAL CRIMINAL LAW 182–87 (3d ed. 2013)).

⁹³ Cf. Kenneth Anderson, Daniel Reisner & Matthew Waxman, *Adapting the Law of Armed Conflict to Autonomous Weapon Systems*, 90 INT’L L. STUD. 386, 410 (2014) (“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.”).

accompanies growing up as a member of a society. Autonomous weapon systems, however, are only as sensible, flexible, and moral as their training data and code—but data sets are notoriously biased,⁹⁴ algorithms are famously inflexible,⁹⁵ and there are myriad examples of artificial intelligence systems acting in surprising ways when programmers incorrectly presumed they would abide by implied, common sense restraints.⁹⁶

Ronald Arkin has suggested that it may be possible to deploy autonomous weapon systems with “Ethical Adaptors,” a metaprogram that would block a “fire” option unless the law of armed conflict and rules of engagement are satisfied.⁹⁷ This top-down approach presumes that a decision tree can be written in advance to address any contingency. Others, concerned that such programming would prove insufficient in the battlefield environment, have proposed supplementary bottom-up approaches, whereby autonomous systems are designed as flexible moral learners.⁹⁸ Both of these proposals assume that technological developments will eventually result in autonomous weapon systems being able to distinguish between a combatant and civilian, between active and wounded or surrendering combatants, and between lawful military objectives and protected civilian objects at least as well as human soldiers. In the meantime, guidelines on appropriate training are needed.

Switching from the weapon to the combatant analogy does not solve the accountability gap. Again, assuming that no one acts intentionally or recklessly, under extant law no one can be held criminally liable if an autonomous weapon commits a serious violation of international humanitarian law.⁹⁹ As of yet,

⁹⁴ See generally, e.g., Kate Crawford, *The Hidden Biases in Big Data*, HARV. BUS. REV. (Apr. 1, 2013), <https://hbr.org/2013/04/the-hidden-biases-in-big-data> [<https://perma.cc/5G7G-MLNC>]; Batya Friedman & Helen Nissenbaum, *Bias in Computer Systems*, 14 ACM TRANSACTIONS ON INFO. SYS. 330 (1996) (discussing preexisting, technical, and emergent bias in computer systems).

⁹⁵ SCHARRE, *supra* note 23, at 6 (“[T]heir brittle nature means that if pushed beyond the bounds of their programming, they may fail, and fail badly.”).

⁹⁶ See, e.g., Victoria Krakovna, *Specification Gaming Examples in AI*, DEEP SAFETY (Apr. 2, 2018), <https://vkrakovna.wordpress.com/2018/04/02/specification-gaming-examples-in-ai/> [<https://perma.cc/3U33-W8HN>] (follow “master list” hyperlink to a collection of examples of unintended behaviors in AI systems); see also Dario Amodei & Jack Clark, *Faulty Reward Functions in the Wild*, OPENAI BLOG (Dec. 21, 2016), <https://blog.openai.com/faulty-reward-functions/> [<https://perma.cc/95GW-K6QJ>] (detailing one example of unexpected AI behavior and discussing how it highlights “a more general issue with reinforcement learning: it is often difficult or infeasible to capture exactly what we want an agent to do, and as a result we frequently end up using imperfect but easily measured proxies. Often this works well, but sometimes it leads to undesired or even dangerous actions.”).

⁹⁷ RONALD C. ARKIN, GOVERNING LETHAL BEHAVIOR IN AUTONOMOUS ROBOTS 138–43 (2009).

⁹⁸ Don Howard & Ioan Muntean, *A Minimalist Model of the Artificial Autonomous Moral Agent (AAMA)*, in 2016 AAAI SPRING SYMPOSIUM SERIES 217, 217–23 (AAAI Publ’ns. ed., 2016), <https://www.aaai.org/ocs/index.php/SSS/SSS16/paper/view/12760/11954>.

⁹⁹ See *supra* note 70 (citing sources).

artificial intelligence has not advanced to a point where an autonomous weapon system could be said to have acted intentionally or recklessly.¹⁰⁰ Nor could it be punished, rendering irrelevant the deterrent aim of individual criminal liability.¹⁰¹ Finally, if an autonomous weapon system is analogized to a human combatant gone rogue, “the deployer could be held directly liable only for actions that resulted in serious violations if she ordered or otherwise directly contributed to the execution of that unlawful action”¹⁰²—which will not be the case when an autonomous weapon system’s act was unforeseen.

Some have argued that command responsibility could be modified to create liability for the actions of autonomous weapon systems, such that a person in putative control could always be held liable.¹⁰³ Doing so, however, would necessitate introducing a criminal negligence standard into international criminal law, which is at odds with the *nulla poena sine culpa* principle, which proscribes that there be no punishment without guilt.¹⁰⁴ As discussed above, the doctrine of command responsibility recognizes the impossibility of completely controlling an autonomous agent’s actions by requiring commanders to punish crimes that they could not prevent. It would be unjust to hold commanders criminally liable for the unpredictable actions of an autonomous weapon system that they can neither predict nor punish.¹⁰⁵

¹⁰⁰ Crootof, *supra* note 46, at 1377 (“If a violation of international humanitarian law is not a war crime absent some willful action, autonomous weapon systems are currently incapable of committing war crimes.”).

¹⁰¹ *But see generally* GABRIEL HALLEVY, WHEN ROBOTS KILL: ARTIFICIAL INTELLIGENCE UNDER CRIMINAL LAW (2013) (arguing that robots can be held criminally liable and punished, much like human beings and corporations); Christina Mulligan, *Revenge Against Robots*, 69 S.C. L. REV. 579 (2018) (arguing that vengeful actions against robotic systems may provide beneficial psychological satisfaction).

¹⁰² Crootof, *supra* note 46, at 1377.

¹⁰³ *See, e.g.*, Christof Heyns (Special Rapporteur on Extrajudicial, Summary, or Arbitrary Executions), *Report of the Special Rapporteur on Extrajudicial, Summary or Arbitrary Executions (Lethal Autonomous Robotics and the Protection of Life)*, ¶ 81, U.N. Doc. A/HRC/23/47 (Apr. 9, 2013) [hereinafter Heyns Report] (“[A]mendments to the rules regarding command responsibility may be needed to cover the use of [autonomous weapon systems].”); Jens David Ohlin, *The Combatant’s Stance: Autonomous Weapons on the Battlefield*, 92 INT’L L. STUD. 1, 26–29 (2016) (arguing for a graduated scheme of criminal liability for commanders who deploy autonomous weapon systems that go on to commit war crimes). Geoffrey Corn has also proposed implementing “procurement responsibility,” under which the military or civilian official who procures a weapon will be responsible for its actions. Geoffrey S. Corn, *Autonomous Weapon Systems: Managing the Inevitability of “Taking the Man Out of the Loop”* 21 (June 14, 2014) (unpublished manuscript), https://papers.ssrn.com/sol3/Delivery.cfm/SSRN_ID2484892_code557106.pdf?abstractid=2450640&mirid=1 [<https://perma.cc/HAU2-E68Y>].

¹⁰⁴ Crootof, *supra* note 46, at 1381–85 (discussing problems with criminalizing negligence).

¹⁰⁵ Instead, I argue that states should be held strictly liable for the actions of their autonomous weapon systems. *See id.* at 1394.

Ultimately, the combatant analogy is largely unsatisfying. It highlights the importance of training independent warfighters in law of armed conflict principles, but provides little guidance on how to train autonomous weapon systems. Nor does it solve the accountability gap problem, as no one will likely be held criminally liable for a latent malfunction that causes harm.

A legal regime for autonomous weapon systems will need to include provisions for legal review, training, and accountability. But the law of weapons does not currently include standards for evaluating systems with the capacity for independent action and in-field learning, the law of combatants does not provide guidance on how to adequately train computer programs, and neither legal regime addresses the accountability gap. Accordingly, neither the weapon nor combatant analogy, standing alone, is sufficient to address the novel legal challenges posed by this new kind of warfighter. But what of other entities that participate in armed conflict?

III. Uncommon Analogies

Autonomous weapon systems, child soldiers, and animal combatants share two legally salient characteristics. They are capable of independent action and, by extension, may sometimes act in unpredictable ways; however, if they take action that resembles a war crime, they cannot be held liable under existing international criminal law. Accordingly, this Section considers child soldiers and animal combatants as potential models for the regulation of autonomous weapon systems.

A. *Child Soldiers*

1. The law of child soldiers

While recruiting or employing child soldiers is a war crime,¹⁰⁶ children continue to play many roles in armed conflicts, including front-line infantry,

¹⁰⁶ The prohibition on the use of child soldiers is articulated in multiple treaties. *See* Optional Protocol to the Convention on the Rights of the Child on the Involvement of Children in Armed Conflict, May 25, 2000, T.I.A.S. No. 13094, 2173 U.N.T.S. 222 [hereinafter *Optional Protocol on Children*] (setting eighteen as the minimum age for direct participation in hostilities and for compulsory recruitment into state armed forces); Convention on the Rights of the Child art. 38, Nov. 20, 1989, 1577 U.N.T.S. 3 (setting fifteen as the minimum age for recruitment or participation in armed conflict); Rome Statute, *supra* note 70, art. 8(2)(b)(xxvi) (including in its definition of war crimes “conscripting or enlisting children under the age of fifteen years” in national armed forces, armed force groups, or using them to participate actively in hostilities); Convention Concerning the Prohibition and Immediate Action for the Elimination of the Worst Forms of Child Labor art. 3(a), June 17, 1999, S. TREATY DOC. NO. 106-5, 2133 U.N.T.S. 161 (including the forced recruitment of individuals under the age of eighteen years for use in armed conflict); First Additional Protocol, *supra* note 8, art. 77(2); Protocol Additional to the Geneva

human mine detectors, suicide bombers, human shields, and sex slaves.¹⁰⁷ Some have been forcibly recruited, but children influenced by various social, economic, and environmental factors also voluntarily enlist.¹⁰⁸ New weapons developments are often celebrated for making war more discriminate and therefore more humane,¹⁰⁹ but ultra-light and precise small arms have also contributed to the usefulness of—and consequently the rise in—child soldiers.¹¹⁰

Given that their participation in armed conflict is itself unlawful, what happens when a child soldier commits a serious violation of international humanitarian law?¹¹¹ As a substantive matter, it is not clear that children have the requisite *mens rea* to commit a war crime. Again, under international law and most domestic law, war crimes must be committed “willfully.”¹¹² But, arguably, before a certain level of development, children are unable to fully understand their acts or the likely consequences, rendering them incapable of acting willfully.¹¹³

As a procedural matter, there is no formal minimum age for international criminal liability. The Geneva Conventions require all member states to punish grave breaches of the Conventions,¹¹⁴ but they do not specify the age of criminal

Conventions of 12 August 1949, and Relating to the Protection of Victims of Non-International Armed Conflicts art. 4(3)(c), June 8, 1977, 1125 U.N.T.S. 609. The prohibition on the recruitment or employment of child soldiers is also likely binding on all states under customary international law. *Rule 136. Recruitment of Child Soldiers*, INT’L COMM. OF THE RED CROSS, https://www.icrc.org/customary-ihl/eng/docs/v1_cha_chapter39_rule136 (last visited May 4, 2018) [<https://perma.cc/Z7G7-PAG7>]; *Rule 137. Participation of Child Soldiers in Hostilities*, INT’L COMM. OF THE RED CROSS, https://www.icrc.org/customary-ihl/eng/docs/v1_cha_chapter39_rule137 (last visited May 4, 2018) [<https://perma.cc/UN43-A2JS>].

¹⁰⁷ See, e.g., Priyanka Boghani, *Why Afghanistan’s Children Are Being Used As Spies and Suicide Bombers*, FRONTLINE (Nov. 17, 2015), <http://www.pbs.org/wgbh/frontline/article/why-afghanistans-children-are-used-as-spies-and-suicide-bombers/> [<https://perma.cc/MSM9-XZ6B>].

¹⁰⁸ Office of the Special Representative of the Secretary-General for Children and Armed Conflict, *Children and Justice During and in the Aftermath of Armed Conflict* 27 (Working Paper No. 3, 2011) [hereinafter *Children in Armed Conflict*], http://reliefweb.int/sites/reliefweb.int/files/resources/Full_Report_1957.pdf [<https://perma.cc/K93Z-UDNH>]; see also *id.* at 28 (noting the Special Representative of the Secretary-General for Children and Armed Conflict’s argument that the distinction between voluntary enlistment and forced recruitment of child soldiers is without meaning, given the often desperate circumstances and environment).

¹⁰⁹ See, e.g., Gross, *supra* note 76, at 48 (discussing how more accurate weaponry allowed for the use of smaller, less-lethal munitions, which in turn reduced the likelihood of collateral damage).

¹¹⁰ See, e.g., Nancy Morisseau, *Seen but Not Heard: Child Soldiers Suing Gun Manufacturers Under Alien Tort Claims Act*, 89 CORNELL L. REV. 1263, 1284–85 (2004).

¹¹¹ For example, Islamic State militants published a video showing child soldiers executing twenty-five men. Lin Jenkins, *ISIS Video Shows Killing of Syrian Troops at Palmyra Amphitheatre*, GUARDIAN (Jul. 4, 2015), <https://www.theguardian.com/world/2015/jul/04/isis-video-killing-palmyra-amphitheatre> [<https://perma.cc/3BM6-WNCC>].

¹¹² See *supra* note 70 (citing sources).

¹¹³ See generally, e.g., Stephen J. Morse, *Immaturity and Irresponsibility*, 88 J. CRIM. L. & CRIMINOLOGY 15 (1997).

¹¹⁴ See, e.g., First Geneva Convention, *supra* note 87, art. 49.

responsibility. There is an argument that fifteen is the minimum age for individual criminal responsibility for war crimes, given that the First Additional Protocol sets fifteen as the minimum age for recruitment and participation.¹¹⁵ If a child is too young to fight, the reasoning goes, the child is too young to be held criminally responsible.¹¹⁶ The International Criminal Court only has jurisdiction to prosecute individuals over the age of eighteen at the time of the alleged crime—not because younger individuals are necessarily exempt from prosecution, but rather because that determination is left to states.¹¹⁷ States, meanwhile, have widely disparate minimum ages for domestic criminal liability, ranging from seven to sixteen.¹¹⁸ So, while an individual under eighteen years of age is unlikely to be tried in an international tribunal for a serious violation of the law of armed conflict, those as young as seven might be prosecuted and convicted in national courts.¹¹⁹ A child soldier's commander, however, may be tried anywhere for the war crime of using child soldiers.¹²⁰

2. The limits of the child soldier analogy

While autonomous weapon systems and child soldiers share the potential for independent action and the inability to act with the requisite *mens rea* to be held individually liable for committing a war crime, child soldiers are not the best analogy for regulatory purposes. The prohibition of child soldiers is grounded primarily on the special rights of children¹²¹ and the resultant obligation to protect them from the physical and psychological dangers associated with armed conflict and the horrors of forced labor and sexual exploitation.¹²² At least at the present stage of technological development, these justifications do not map well onto

¹¹⁵ Children in Armed Conflict, *supra* note 108, at 34.

¹¹⁶ *Id.*

¹¹⁷ Rome Statute, *supra* note 70, art. 26. Radhika Coomaraswamy, then-Special Representative for Children and Armed Conflict, has noted that there are problematic incentives for setting the minimum age for criminal liability too high: “If minor children who have committed serious war crimes are not prosecuted, this could be an incentive for their commanders to delegate to them the dirtiest orders, aiming at impunity.” Radhika Coomaraswamy, *The Optional Protocol to the Convention on the Rights of the Child on the Involvement of Children in Armed Conflict—Towards Universal Ratification*, 18 INT’L J. CHILDREN’S RTS. 535, 544 (2010).

¹¹⁸ Children in Armed Conflict, *supra* note 108, at 35.

¹¹⁹ Conviction need not imply capital punishment or even imprisonment. *Cf.* Convention on the Rights of the Child, *supra* note 106, art. 37(a). Instead, “the purpose of any sanction imposed on a child should be to promote rehabilitation and reintegration in to the community and not to punish.” *See* Children in Armed Conflict, *supra* note 108, at 36. Indeed, holding former child soldiers accountable in a way that fosters reintegration may be necessary to their personal recovery.

¹²⁰ *See* Prosecutor v. Dyilo, Case No. ICC-01/04-01/06-2842, Judgment (Mar. 14, 2012) (finding defendant guilty of recruiting and using child soldiers); Prosecutor v. Taylor, Case No. SCSL-03-01-T, Judgment (Spec. Ct. for Sierra Leone May 18, 2012) (finding defendant guilty of supporting the recruitment and use of child soldiers).

¹²¹ *See, e.g.*, Convention on the Rights of the Child, *supra* note 106.

¹²² *See, e.g.*, Optional Protocol on Children, *supra* note 106, pmb1.

autonomous weapon systems: there is no similar need to protect them from the horrors of war.¹²³

It is worth noting that, because the use of child soldiers is itself unlawful, incurring both individual criminal liability and state responsibility,¹²⁴ there is no accountability gap. Commanders employing child soldiers are liable for the war crime of using child soldiers—by extension, they should also be liable for any unlawful actions of child soldiers they recruit or supervise.¹²⁵ But there is no corresponding ban on the use of autonomous weapon systems, nor is there likely to be.¹²⁶

Given the fundamental differences between the reasons for regulating the use of child soldiers and autonomous weapon systems, this analogy is ultimately unhelpful.

B. *Animal Combatants*

1. The law of animal combatants

Animals have long participated in armed conflicts: “The variety of animal species deployed as ‘soldiers’ in armed conflicts ranges from horses, elephants, dogs, bats, camels, seals and pigeons to dolphins, bees, donkeys, belugas, oxen and cormorants.”¹²⁷ Animals have served with distinction—military medals and

¹²³ *But see* Kate Darling, *Extending Legal Protection to Social Robots: The Effects of Anthropomorphism, Empathy, and Violent Behavior Towards Robotic Objects*, in *ROBOT LAW*, *supra* note 5, 213, 213–32 (arguing that, as was the underlying purpose of animal abuse laws, we should grant social robots legal protections from “harm” precisely because we anthropomorphize them); Christopher D. Stone, *Should Trees Have Standing?—Toward Legal Rights for Natural Objects*, 45 S. CAL. L. REV. 450, 453–55 (1972) (observing that, “each time there is a movement to confer rights onto some new ‘entity,’ the proposal is bound to sound odd or frightening or laughable” and that many entities once considered “things” under law—including African Americans, fetuses, and women—are now independent rights-holders).

¹²⁴ *See* sources cited *supra* note 106 and accompanying text.

¹²⁵ Children in Armed Conflict, *supra* note 108, at 37. The Chief Prosecutor of the Special Court of Sierra Leone has stated that he would not indict persons for crimes committed when they were children, but instead would indict their adult recruiters and commanders. Press Release, Special Ct. for Sierra Leone Pub. Aff. Off., Special Court Prosecutor Says He Will Not Prosecute Children (Nov. 2, 2002), <http://www.rscsl.org/Documents/Press/OTP/prosecutor-110202.pdf> [<https://perma.cc/AS9P-CMEG>].

¹²⁶ Autonomous weapon systems share few qualities with weapons that have been successfully banned, and states are unlikely to voluntarily relinquish their right to use this uniquely effective, incompletely understood, and multifaceted tool. Crotoft, *supra* note 2, at 1891–93; Sean Watts, *Autonomous Weapons: Regulation Tolerant or Regulation Resistant?*, 30 *TEMPLE INT’L & COMP. L.J.* 177, 178 (2016).

¹²⁷ Karsten Nowrot, *Animals at War: The Status of ‘Animal Soldiers’ Under International Humanitarian Law*, 40 *HIST. SOC. RES.* 128, 130 (2015) (citation omitted).

decorations have been awarded to horses, dogs, pigeons, and cats¹²⁸—and they have been formally detained. In February 2014, the Afghanistan Taliban revealed that they had captured a British military working dog, and a Taliban spokesman offered assurances that the dog “was not injured and is not being mistreated.”¹²⁹ Dogs in particular have played many roles in armed conflict, from sentries to transporters to bomb detectors.¹³⁰ In World War II, Russian dogs were trained to carry explosives under German tanks in suicide missions.¹³¹ Today, states have various animal combatant divisions.¹³²

However, animal combatants have no formal status or protections under international humanitarian law.¹³³ The Geneva Conventions presume that only human beings can be members of an armed force, militia, or volunteer corps;¹³⁴ indeed, the single international humanitarian law treaty reference to animals in armed conflict is in Protocol II of the Convention on Certain Conventional Weapons, which bans the use of booby-traps or other devices that are attached to or associated with “animals or their carcasses.”¹³⁵

What, then, would happen if an animal combatant were to take an action that resulted in what seemed to be a serious violation of international humanitarian law? Imagine if a war dog accidentally carried an explosive under a

¹²⁸ *Id.* at 132.

¹²⁹ Amir Shah, *Taliban: Dog Now a POW in Afghanistan*, CHRISTIAN SCI. MONITOR (Feb. 8, 2014), <http://www.csmonitor.com/World/Latest-News-Wires/2014/0208/Taliban-Dog-now-a-POW-in-Afghanistan>.

¹³⁰ *Types of War Dogs*, U.S. WAR DOG ASS'N, <http://www.uswardogs.org/war-dog-history/types-war-dogs/> (last visited May 4, 2018) [https://perma.cc/NPX8-9LV5].

¹³¹ Although there are individual examples of successes, this approach was largely deemed inefficient. One serious mistake was that Russians trained the dogs on their own diesel-engine tanks, rather than on German gasoline-engine tanks. As a result, armed dogs would seek out the familiar-smelling Russian tanks on the battlefield. See CHRIS BISHOP, *THE ENCYCLOPEDIA OF WEAPONS OF WORLD WAR II* 205 (2002).

¹³² See, e.g., Military Working Dogs, Army Reg. 190-12 (2013) (discussing the Military Working Dogs Program); *Marine Mammal Program*, U.S. NAVY, <http://www.public.navy.mil/spawar/Pacific/technology/Pages/mammals.aspx> (last visited May 4, 2018) [https://perma.cc/8C5W-SPL3].

¹³³ Nowrot, *supra* note 127, at 135 (“[D]espite the fact that certain animals are quite frequently allowed or required to ‘participate directly in hostilities’ . . . they are not granted the rights and do not have the obligations deriving from the legal status of combatants under international humanitarian law.”).

¹³⁴ Cf. Third Geneva Convention, *supra* note 78, art. 4(a) (limiting prisoners of war to “persons” belonging to various categories).

¹³⁵ Protocol on Prohibitions or Restrictions on the Use of Mines, Booby-Traps and Other Devices (Protocol II) art. 7(1), Oct. 10, 1980, S. TREATY DOC. No. 105-1, 2048 U.N.T.S. 133 (amended May 3, 1996). Animals that are not associated with military forces are protected generally as civilian objects. See, e.g., First Additional Protocol, *supra* note 8, art. 54 (prohibiting the destruction and removal of livestock as “objects indispensable to the survival of the civilian population”).

civilian school bus or ambulance, resulting in civilian death. In the Middle Ages, it was not unheard of to hold animals accountable for their tortious and criminal actions.¹³⁶ A dog was imprisoned for a year;¹³⁷ a milk cow was killed for murdering a pregnant woman,¹³⁸ more than once, pigs were condemned to death for infanticide.¹³⁹ Today, these practices seem ludicrous. But it is unclear whether any other entity would be held responsible for the animal's actions.¹⁴⁰

2. The limits of the animal combatant analogy

In many ways, animals are the best analogy for autonomous weapon systems. They are not quite weapons, insofar as they are capable of independent and unpredictable action. They are not quite combatants: they cannot be taught the law of armed conflict, and they cannot act with the requisite *mens rea* for criminal liability. Their independence is tempered through extensive training; their propensity for unpredictable action is addressed through limited use.

While the animal analogy is conceptually useful,¹⁴¹ there is no international “law of animal combatants.” The lack of explicit prohibitions or regulations suggests that sovereign states may do what they wish, providing little guidance regarding appropriate international regulatory standards for autonomous weapon systems.¹⁴²

¹³⁶ See William Ewald, *Comparative Jurisprudence (I): What Was It Like To Try a Rat?*, 143 U. PA. L. REV. 1889 (1995) (discussing the animal trials of the Middle Ages). According to Ewald, “[f]rom the ninth century to the nineteenth, in Western Europe, there are over two hundred well-recorded cases of trials of animals.” *Id.* at 1903. Nor did human plaintiffs always prevail: suits against animals were dismissed, sometimes on procedural and sometimes on substantive grounds. *Id.* at 1898–99 (discussing the 1522 trial of the rats for having eaten and wantonly destroyed barley crops—when the rats failed to appear in court after the first and second summons, their advocate successfully argued that they had not received adequate notice); *id.* 1902–03 (discussing the 1545 and 1587 trials of snout-beetles for infesting vineyards—the original suit was dismissed on the grounds that God had provided sustenance for all of his creatures).

¹³⁷ *Id.* at 1905.

¹³⁸ *Id.* at 1904.

¹³⁹ *Id.*

¹⁴⁰ *But see* Nowrot, *supra* note 127, at 142 (suggesting that human handlers would be held accountable for the acts of animal combatants).

¹⁴¹ Heather Roff and David Danks discuss the pros and cons of incorporating “AWS liaisons” in military forces, building off of existing practices regarding animals and human handlers. Heather M. Roff & David Danks, *The Necessity and Limits of Trust in Autonomous Weapon Systems*, 17 J. MIL. ETHICS (forthcoming 2018).

¹⁴² The *Lotus* court established the principle that state actions not expressly prohibited under international law are permitted. S.S. *Lotus* (Fr. v. Turk.), Judgment, 1927 P.C.I.J. (ser. A) No. 10, at 18 (Sept. 7) (“Restrictions upon the independence of States cannot . . . be presumed.”). *But see* Unilateral Declaration of Independence in Respect of Kosovo, Advisory Opinion, 2010 I.C.J. Rep. 404 ¶ 8 (July 22) (declaration of Simma. J.) (criticizing the *Lotus* principle as an outdated and “excessively deferential” approach).

IV. The Limits of Analogy

Many have argued that existing international humanitarian law can address the regulatory challenges raised by autonomous weapon systems,¹⁴³ in part because most autonomous weapon systems in use today can be analogized to other weapons. But the analogy that works today will not necessarily work tomorrow. Weapon systems with greater levels of autonomy and in-field learning capabilities are being developed, prompting the question of how long the weapon analogy will remain accurate or useful. Meanwhile, the combatant, child soldier, and animal combatant analogies misrepresent other important characteristics. Collectively, these analogies constrain how we think about autonomous weapon systems, thereby impeding our ability to craft appropriate regulations.

A. *Analogies Are Misleading*

All of the aforementioned analogies misrepresent legally salient traits of autonomous weapon systems: the weapon analogy minimizes their capacity for independent action, while the combatant, child soldier, and animal analogies overemphasize it.

As with definitions that would lump landmines and Skynet in the same category,¹⁴⁴ the weapon analogy “almost certainly [misses] the essence of what is new about autonomous weapons.”¹⁴⁵ Autonomous weapon systems are fundamentally different from inert, automated, and semi-autonomous weapons because of their ability to independently select and engage targets. In addition to the moral questions raised by partially delegating life and death decisions to algorithms, this capacity raises legal issues that simply are not addressed by the existing law governing weapons.

¹⁴³ Charlie Dunlap, for example, has long argued that “the best way to regulate any weapon (to include autonomous and other high-tech weapons) is by insisting that it strictly adhere to the *existing* law of war (as opposed to trying to create a specialized legal regime for every new technology that appears).” Charlie Dunlap, *Autonomous Weapons and the Law: The Yale and Brookings Discussions*, LAWFIRE (Apr. 9, 2018), <https://sites.duke.edu/lawfire/2018/04/09/autonomous-weapons-and-the-law-the-yale-and-brookings-discussions/> [https://perma.cc/X8ZH-BWXX]; see also Charlie Dunlap, *A Better Way to Protect Civilians and Combatants Than Weapons Bans: Strict Adherence to the Core Principles of the Law of War*, LAWFIRE (Dec. 3, 2015), <https://sites.duke.edu/lawfire/2015/12/03/a-better-way-to-protect-civilians-and-combatants-than-weapons-bans-strict-adherence-to-the-core-principles-of-the-law-of-war-2/> [https://perma.cc/2LK6-ANVR].

¹⁴⁴ See Crotoof, *supra* note 2, at 1851–52; THE TERMINATOR, *supra* note 3 (including Skynet, a self-aware artificial superintelligence).

¹⁴⁵ Paul Scharre, *Autonomy, “Killer Robots,” and Human Control in the Use of Force—Part I*, JUST SECURITY (July 9, 2014), <http://justsecurity.org/12708/autonomy-killer-robots-human-control-force-part/> [https://perma.cc/2P3H-KLC9].

Meanwhile, the combatant, child soldier, and animal analogies inappropriately characterize autonomous weapon systems as fully independent entities. Unlike combatants, child soldiers, and animals—who may be coerced or tricked but nevertheless retain a fundamental degree of autonomy—autonomous weapon systems may have their ability to act independently sharply curtailed. The actions of autonomous weapon systems are based on algorithms, and algorithms may be rewritten. Autonomous weapon systems may be operated solely in semi-autonomous or even automatic modes by their deployers, or they may be hacked and completely controlled by an enemy.¹⁴⁶

B. Analogies Are Constraining

Analogies are also inherently constraining, in that they restrict our ability to think imaginatively about a new technology. Consider the use of the term “driverless cars” to describe autonomous vehicles. The inherent analogy normalizes something new and dangerous, but it also restricts our understanding and imagination. There is no reason to think autonomous vehicles will look or operate anything like existing cars, just as early cars did not look or operate like “horseless carriages.” An autonomous vehicle need not have a steering wheel or other means of human interaction with the system. And conceiving of autonomous vehicles as driverless cars locks one into a host of existing assumptions, instead of allowing for more imaginative conceptions of what the technology might permit. For example, rather than being individually owned and operated property, autonomous vehicles could operate as connected nodes on a “smart highway” or as a leasable service.¹⁴⁷ Similarly, thinking of autonomous weapon systems as a single, independent, embodied entity—be it a weapon, combatant, child, or animal—prevents us from imagining what other forms they might take.

One possibility is that autonomous weapon systems will be a collection of networked systems. The U.S. Navy’s LOCUST (low-cost UAV swarming technology) can already launch up to thirty small drones that communicate with each other to fly in formation and engage in “defensive or offensive missions.”¹⁴⁸ Now imagine vast systems of widespread sensors; unmanned aerial, underwater,

¹⁴⁶ SCHARRE, *supra* note 23, at 14–15. Alternatively, autonomous weapon systems may have their capabilities dramatically expanded. Autonomous weapon systems can “learn” through millions of simulated experiences, they can download information learned by other systems, and they can be augmented with new software and hardware.

¹⁴⁷ For a fascinating analysis of the different forms autonomous vehicles might take and their various legal implications, see Jack Boeglin, *The Costs of Self-Driving Cars: Reconciling Freedom and Privacy with Tort Liability in Autonomous Vehicle Regulation*, 17 YALE J.L. & TECH. 171 (2015).

¹⁴⁸ Press Release, David Smalley, Off. of Naval Res., LOCUST: Autonomous, Swarming UAVs Fly into the Future (Apr. 14, 2015) <https://www.onr.navy.mil/Media-Center/Press-Releases/2015/LOCUST-low-cost-UAV-swarm-ONR> [<https://perma.cc/HW97-964P>].

or surface vehicles; and central “brains”: each component, individually, might not constitute an autonomous weapon nor present much of a challenge for traditional legal reviews of new weaponry, but collectively this system would enable an entirely new means of waging war.

Alternatively, autonomous weapon systems might take the form of “centaur corps,” human-machine teams designed to leverage the strengths of both entities.¹⁴⁹ An AI system named ALPHA made headlines for besting a retired U.S. Air Force colonel in multiple flight simulator trials, but it was designed to assist, rather than replace, human pilots, either by providing real-time advice or by flying protective “wingmen” UAVs.¹⁵⁰ States also continue to invest in research and development of powered, armored exoskeleton suits like the U.S. Special Operations Command’s Tactical Assault Light Operator Suit (TALOS), designed with power-assisted limbs, 360-degree night vision sensors, and open architecture that will allow for a variety of add-on capabilities.¹⁵¹ Accordingly, some have argued that it may be necessary to expand the legal review for new weapons to encompass augmented or cyborg combatants.¹⁵²

Autonomous weapon systems need not be embodied at all: autonomous cyberweapons are already widely deployed,¹⁵³ and the speed of cyber will nearly always require that countermeasures be automated or autonomous to be effective.¹⁵⁴ And while physical autonomous weapon systems used for defensive purposes are relatively uncontroversial—indeed, ban advocates usually exempt

¹⁴⁹ See SCHARRE, *supra* note 23, at 5 (“‘Centaur’ human-machine teaming cognitive architectures can leverage the predictability, reliability, and speed of automation while retaining the robustness and flexibility of human intelligence.”). This concept is grounded on “centaur chess,” in which human-machine teams have proven superior to either human or machine chess players. *Id.* at 39–40.

¹⁵⁰ Clint Finley, *AI Fighter Pilot Beats a Human, But No Need to Panic (Really)*, WIRED (Jun. 29, 2016), http://www.wired.com/2016/06/ai-fighter-pilot-beats-human-no-need-panic-really/?mbid=social_fb [https://perma.cc/AKG7-4VPW].

¹⁵¹ Sean Gallagher, *DOD Continues Quest to Make “Iron Man” Exosuit for Special Ops*, ARS TECHNICA (Jun. 2, 2016), <http://arstechnica.com/information-technology/2016/06/dod-continues-quest-to-make-iron-man-exosuit-for-special-ops/> [https://perma.cc/69UA-8DMC]. For a general discussion of military, maintenance, medical, and other uses of exosuits in maritime operations, see ANDREW HERR & LT. SCOTT CHENEY-PETERS, BETWEEN IRON MAN AND AQUA MAN: EXOSUIT OPPORTUNITIES IN MARITIME OPERATIONS (2015), https://s3.amazonaws.com/files.cnas.org/documents/CNAS_IronManAndAquaMan_HerrCheneyPeters.pdf?mtime=20160906081329 [https://perma.cc/P26W-THT9].

¹⁵² Patrick Lin, *Could Human Enhancement Turn Soldiers Into Weapons That Violate International Law? Yes*, ATLANTIC, Jan. 4, 2013, <http://www.theatlantic.com/technology/archive/2013/01/could-human-enhancement-turn-soldiers-into-weapons-that-violate-international-law-yes/266732/> [https://perma.cc/EM2Z-WKYB].

¹⁵³ See generally VIGNARD, *supra* note 25.

¹⁵⁴ See Eric Messinger, *Is It Possible to Ban Autonomous Weapons in Cyberwar?*, JUST SECURITY (Jan. 15, 2015), <https://www.justsecurity.org/19119/ban-autonomous-weapons-cyberwar/> [https://perma.cc/D27R-2425].

such systems from their definition of autonomous weapon systems altogether¹⁵⁵—autonomous defensive cyberweapons will likely encourage states to develop autonomous offensive cyberweapons.¹⁵⁶ Furthermore, there is “an important dimension to malware that does not have an analogy in physical weapons”: the ease with which cyberweapons can self-replicate and spread.¹⁵⁷

By misrepresenting legally salient traits and by limiting our ability to imagine future developments, analogies can sometimes impede our ability to appropriately regulate new technology. Instead of borrowing from different legal regimes and relying on analogical reasoning to create a patchwork of regulations for autonomous weapon systems,¹⁵⁸ we should instead acknowledge the fundamentally distinct nature of these new warfighters and create appropriate, supplemental law.

Conclusion

While analogical reasoning allows “*most* law-of-war rules [to] apply *most* of the time to *most* new technologies,”¹⁵⁹ in some situations there is no way to credibly stretch existing rules to answer novel legal questions. Autonomous weapon systems raise a host of issues that cannot be addressed by analogical reasoning, as all potential analogies misrepresent legally salient traits. Conceiving of autonomous weapon systems merely as weapons minimizes their capacity for

¹⁵⁵ For example, in July 2015 the Future of Life Institute published an Open Letter from AI and Robotics Researchers calling for a ban on “offensive autonomous weapon systems beyond meaningful human control.” *Autonomous Weapons: An Open Letter from AI & Robotics Researchers*, FUTURE OF LIFE INST. (July 28, 2015), <https://futureoflife.org/open-letter-autonomous-weapons/> [<https://perma.cc/HYY3-JB8D>]. *But see* Crotoft, *supra* note 2, at 1863 (arguing that whether a weapon system is being used for offensive or defensive purposes is a question of use, not autonomy).

¹⁵⁶ Messinger, *supra* note 154. The Defense Advanced Research Projects Agency has recently hosted a cyber challenge, wherein it invited hacker teams to create programs that can independently identify and patch security flaws in a system. *The World’s First All-Machine Hacking Tournament: Event Recap*, DEF. ADVANCED RES. PROJECTS AGENCY, <http://archive.darpa.mil/cybergrandchallenge/> (last visited Mar. 27, 2018) [<https://perma.cc/DC6V-NJ4K>] (recapping DARPA’s Cyber Grand Challenge). A similarly autonomous program might be used to identify and exploit enemy vulnerabilities.

¹⁵⁷ SCHARRE, *supra* note 25, at 225 (“Drones and robotic systems cannot self-replicate. In this sense, malware resembles biological viruses and bacteria, which self-replicate and spread host to host.”).

¹⁵⁸ For a review of a few of the legal regimes that may be relevant to how states may lawfully develop or employ autonomous weapon systems—including international human rights law, the law of the sea, space law, and the law of state responsibility—see Rebecca Crotoft, *The Varied Law of Autonomous Weapon Systems*, in NATO ALLIED COMMAND TRANSFORMATION, AUTONOMOUS SYSTEMS: ISSUES FOR DEFENCE POLICY MAKERS 98, 98–126 (Andrew P. Williams & Paul D. Scharre eds., 2015).

¹⁵⁹ Kristen E. Eichensehr, *Cyberwar and International Law Step Zero*, 50 TEX. INT’L L.J. 357, 359 (2015).

independent and self-determined action, while the combatant, child soldier, and animal combatant comparisons overemphasize it. Furthermore, all of these discrete and embodied analogies limit our ability to think imaginatively about this new technology and anticipate how it might develop, thereby impeding our ability to properly regulate it.

We cannot simply graft legal regimes designed to regulate weapons or human combatants onto autonomous weapon systems. Instead, as is often the case when there is no appropriate analogy, it is time to explicitly revise rules or create entirely new ones to address the specific situations where extant law is insufficient. We need to develop standards for both the training and the legal review of increasingly independent autonomous weapon systems,¹⁶⁰ and we should outline accountability mechanisms for their generally inevitable but individually unforeseeable accidents.¹⁶¹ More broadly, we should be having a larger conversation about the amount of human oversight, control, or judgment necessary in the targeting process for all attacks.

The sooner we acknowledge the insufficiency of existing analogies, the sooner we can create appropriate and effective regulations for autonomous weapon systems.

¹⁶⁰ Farrant & Ford, *supra* note 24, at 406–07 (discussing Article 36 legal reviews for continuously learning systems).

¹⁶¹ Crootof, *supra* note 46, at 1386 (arguing that states should be held strictly liable for the harms caused by their autonomous weapon systems).