

2016

## A Meaningful Floor for "Meaningful Human Control"

Rebecca Crootof

*University of Richmond - School of Law*, [rcrootof@richmond.edu](mailto: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, *A Meaningful Floor for "Meaningful Human Control"*, 30 Temp. Int'l & Comp. L.J. 53 (2016).

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](mailto:scholarshiprepository@richmond.edu).

# A MEANINGFUL FLOOR FOR “MEANINGFUL HUMAN CONTROL”

*Rebecca Crootof* \*

## I. INTRODUCTION

To the extent there is any consensus among States, ban advocates, and ban skeptics regarding the regulation of autonomous weapon systems (AWS),<sup>1</sup> it is grounded in the idea that all weaponry should be subject to “meaningful human control.”<sup>2</sup> This “intuitively appealing” principle is immensely popular,<sup>3</sup> and numerous States have explicitly declared their support for it or questioned the lawfulness of weapons that operate without such control.<sup>4</sup> Lack of opposition has

---

\* Executive Director, Yale Information Society Project; Research Scholar and Lecturer in Law, Yale Law School. Thanks to Duncan Hollis and the other organizers of *Autonomous Legal Reasoning? Legal and Ethical Issues in the Technologies of Conflict* for the opportunity to explore this issue; thanks to Kenneth Anderson and other symposium participants for helpful commentary on earlier drafts.

1. 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 (2015); see U.S. DEP’T OF DEF., DIRECTIVE NO. 3000.09: AUTONOMY IN WEAPON SYSTEMS 13–14 (2012); INT’L COMM. OF THE RED CROSS, INTERNATIONAL HUMANITARIAN LAW AND THE CHALLENGES OF CONTEMPORARY ARMED CONFLICTS 44 (2015), <https://www.icrc.org/en/document/international-humanitarian-law-and-challenges-contemporary-armed-conflicts> [hereinafter ICRC REPORT] (“[A]utonomous weapon systems’ is an umbrella term that would encompass any type of weapon systems, whether operating in the air, on land or at sea, with autonomy in its ‘critical functions,’ meaning a weapon that can select . . . and attack . . . targets without human intervention.”).

2. Michael Biontino (Chairperson of the Informal Meeting of Experts), *Report of the 2015 Informal Meeting of Experts on Lethal Autonomous Weapons Systems (LAWS)*, UN OFF. GENEVA (2015) at 11, <http://reachingcriticalwill.org/images/documents/Disarmament-fora/ccw/2015/DraftReport.pdf> [hereinafter Biontino Report] (“There seems to be a widespread understanding that both the legal and ethical acceptability of a weapon system would require some kind of human control.”).

3. U.N. INST. FOR DISARMAMENT RESEARCH, THE WEAPONIZATION OF INCREASINGLY AUTONOMOUS TECHNOLOGIES: CONSIDERING HOW MEANINGFUL HUMAN CONTROL MIGHT MOVE THE DISCUSSION FORWARD 2 (2015) [hereinafter UNIDIR REPORT].

4. See, e.g., CAMPAIGN TO STOP KILLER ROBOTS, COUNTRY POLICY POSITIONS (2015), [http://www.stopkillerrobots.org/wp-content/uploads/2015/03/KRC\\_CCWexperts\\_Countries\\_25Mar2015.pdf](http://www.stopkillerrobots.org/wp-content/uploads/2015/03/KRC_CCWexperts_Countries_25Mar2015.pdf) (compiling statements from States regarding meaningful human control). As this paper was being finalized for publication, the 2016 CCW Meeting of Experts on Lethal Autonomous Weapon Systems was concluding. As in past years, States voiced support for maintaining meaningful human control over weapons; and, as in past years, what that entailed remained contested. Chris Ford & Chris Jenks, *The International Discussion Continues: 2016 CCW Experts Meeting on Lethal Autonomous Weapons*, JUST SEC. (Apr. 20, 2016), <https://www.justsecurity.org/30682/2016-ccw-experts-meeting-laws/>; Kelley Saylor, *More of the*

led some to conclude that it is either a newly developed customary norm or a pre-existing, recently exposed rule of customary international law, already binding on all States.<sup>5</sup>

But this broad support comes at a familiar legislative cost; there is no consensus as to what “meaningful human control” actually requires. State X might define meaningful human control to require informed human approval of each possible action of a given weapon system (maintaining a human being “in the loop”);<sup>6</sup> State Y might understand it as the ability of a human operator to oversee and veto a weapon system’s actions (having a human being “on the loop”); and State Z might view the original programming alone as providing sufficiently meaningful human control (allowing human beings to be “off the loop”).<sup>7</sup> As the Czech Republic noted, in voicing its belief that “the decision to end somebody’s life must remain under meaningful human control, . . . [t]he challenging part is to establish what precisely ‘meaningful human control’ would entail.”<sup>8</sup>

This paper describes attempts to clarify what factors are relevant to meaningful human control, discusses benefits associated with retaining imprecision in a standard intended to regulate new technology through international consensus, and argues that the standard’s vagueness should be limited

*Same: The UN Debate on Lethal Robots*, CTR. FOR THE STUDY OF THE DRONE (Apr. 27, 2016), <http://dronecenter.bard.edu/more-of-the-same-the-un-debate-on-lethal-robots/>.

5. Peter Asaro, Jus Nascendi, *Robotic Weapons and the Martens Clause*, in *ROBOT LAW* (Ryan Calo, Michael Froomkin & Ian Kerr eds.) (forthcoming 2016) (on file with author) (“[T]here is already an emerging norm concerning meaningful human control over the targeting of weapons and the use of violent force . . . .”); see also Michael C. Horowitz & Paul Scharre, *Meaningful Human Control in Weapon Systems: A Primer* 7 (Ctr. for New Am. Sec., Working Paper No. 031315, 2015), [http://www.cnas.org/sites/default/files/publications-pdf/Ethical\\_Autonomy\\_Working\\_Paper\\_031315.pdf](http://www.cnas.org/sites/default/files/publications-pdf/Ethical_Autonomy_Working_Paper_031315.pdf) (describing two main schools of thought regarding the principle’s legal status).

6. This understanding is favored by pro-ban advocates as it essentially permits only the development and use of semi-autonomous weapon systems. See Thomas Nash, Director, Article 36, Remarks to the CCW on Autonomous Weapons Systems (May 15, 2014), <http://www.article36.org/statements/701/> (“[M]eaningful human control . . . can be seen to require deliberative moral reasoning, by human beings, over individual attacks. Weapons that do not allow such human control and attacks without such human control should be prohibited.”); Kelsey D. Atherton, *The International Community is About to Debate Killer Robots*, *POPULAR SCI.* (Apr. 11, 2016), <http://www.popsoci.com/international-community-is-about-to-debate-killer-robots> (quoting Mark Guburd: “If a (non-human) system makes a decision under internal programming plus environmental inputs, that is not human control. . . . You may have programmed it, and you may be satisfied that it is making the right decisions, but you are not controlling it when it makes those decisions. . . . The whole point of calling something autonomous is that it is operating outside of human control, making decisions on its own.”).

7. See Charli Carpenter, *Dynamics of Debate at the Experts Meeting on Autonomous Weapons*, *DUCK OF MINERVA* (May 15, 2014), <http://duckofminerva.com/2014/05/dynamics-of-debate-at-the-experts-meeting-on-autonomous-weapons.html> (“A U.S. delegate implied yesterday that autonomous weapons would of course be controlled by humans because as the creators of these weapons we are responsible for their programming.”).

8. Statement of the Czech Republic, CCW Meeting of Experts on Lethal Autonomous Weapons Systems, Geneva, April 13–17, 2015, [http://www.unog.ch/80256EDD006B8954/\(httpAssets\)/2DD5110A33C9C2D2C1257E26005DD47B/\\$file/2015\\_LAWS\\_MX\\_Czech+Republic.pdf](http://www.unog.ch/80256EDD006B8954/(httpAssets)/2DD5110A33C9C2D2C1257E26005DD47B/$file/2015_LAWS_MX_Czech+Republic.pdf).

by an interpretive floor. “Meaningful human control” as a regulatory concept can usefully augment existing humanitarian norms governing targeting—namely, that all attacks meet the treaty and customary international law requirements of distinction, proportionality, and feasible precautions.<sup>9</sup> However, it should not be interpreted to conflict with these norms nor be prioritized in a way that undermines existing humanitarian protections.

## II. INHERENT IMPRECISION

The phrase “meaningful human control” was first influentially used in a briefing paper to the 2014 Experts Meeting on Lethal Autonomous Weapon Systems.<sup>10</sup> The paper raised important questions that still warrant discussion: what type of human control should be exercised over autonomous weapon systems? At what point is that control no longer meaningful? To what extent can computer programming augment or even constitute human control?<sup>11</sup>

In part because it is so difficult to pin down precisely what is uniquely concerning about AWS<sup>12</sup>—especially given that so many automatic, semi-autonomous, and even quietly autonomous weapons are in use today<sup>13</sup>—this phrase quickly captured the collective imagination.<sup>14</sup> Pro-ban advocates have even begun to define AWS as those that operate “without” or “beyond” meaningful human control.<sup>15</sup> But there are nearly as many understandings of what meaningful human

---

9. Cf. Michael N. Schmitt, *Military Necessity and Humanity in International Humanitarian Law: Preserving the Delicate Balance*, 50 *VIR. J. INT’L L.* 795 (2010) (discussing these requirements and arguing that there is no need to consider an independent “military necessity” or “humanity” requirement, as both are implicitly considered and balanced in the other requirements).

10. *Key Areas for Debate on Autonomous Weapon Systems*, ARTICLE 36 (May 2014), <http://www.article36.org/wp-content/uploads/2014/05/A36-CCW-May-2014.pdf>.

11. *Id.* at 2.

12. See generally Michael C. Horowitz, *The Ethics & Morality of Robotic Warfare: Assessing the Debate Over Autonomous Weapons*, 145 *J. OF AM. ACAD. OF ARTS & SCI.* 25 (Fall 2016).

13. See Crotoft, *supra* note 1, at 1863–72 (describing various autonomous weapons systems in use today); see also *Autonomous Weapons: Decisions to Kill and Destroy Are A Human Responsibility*, INT’L COMM. RED CROSS (Apr. 2016), <https://www.icrc.org/en/document/statement-icrc-lethal-autonomous-weapons-systems> [hereinafter ICRC Statement] (“Some weapon systems in use today can select and attack targets without human intervention.”).

14. See, e.g., ICRC Statement, *supra* note 13 (“Whether for legal, ethical or military-operational reasons, there is broad agreement on the need for human control over weapons and the use of force.”).

15. See HUMAN RIGHTS WATCH & INT’L HUMAN RIGHTS CLINIC, HARVARD LAW SCH., *MIND THE GAP: THE LACK OF ACCOUNTABILITY FOR KILLER ROBOTS 1* (2015), [https://www.hrw.org/sites/default/files/reports/arms0415\\_ForUpload\\_0.pdf](https://www.hrw.org/sites/default/files/reports/arms0415_ForUpload_0.pdf) [hereinafter MIND THE GAP] (“Fully autonomous weapons, also known as ‘killer robots,’ raise serious moral and legal concerns because they would possess the ability to select and engage their targets without meaningful human control.”); *The United Kingdom and Lethal Autonomous Weapon Systems*, ARTICLE 36, at 1 (2016), <http://www.article36.org/wp-content/uploads/2016/04/UK-and-LAWS.pdf> (“The defining feature of lethal autonomous weapons systems is that they would be systems that operate without meaningful human control.”).

control entails as there are writers on the subject.<sup>16</sup>

Certainly, there are identifiable situations at the far poles of the spectrum. When a well-trained and informed person monitors and approves every action taken by a weapon system, there is little doubt that she is exercising meaningful human control. When a poorly trained, uninformed person “robotically” pushes a button every time a red light goes on, she is merely a human rubber stamp—and likely a legal and moral scapegoat.<sup>17</sup>

But the grey area is wide, and full of complicated situations. For example, the speed of cyber actions makes automatic and even autonomous cyber defenses increasingly desirable.<sup>18</sup> As a result, “in offensive cyberwarfare, [autonomous weapon systems] may *have* to be deployed, because they will be integral to effective action in an environment populated by automated defenses and taking place at speeds beyond human capacities.”<sup>19</sup> In both cyber defense and offense, requiring in-time human approval for every action would be a dangerous impediment. What should the standard for meaningful human control be in cyberspace? And should it differ from the standard governing weaponry in other realms?

There is a growing scholarly literature attempting to clarify what factors are relevant to establishing meaningful human control.<sup>20</sup> In early 2014, the International Committee for Robot Arms Control (ICRAC) proffered three “minimum necessary conditions for meaningful [human] control”<sup>21</sup>:

---

16. See ICRC Statement, *supra* note 13 (noting the need for clarity about what “kind and degree of human control” should be exercised, ideally in a way that satisfies legal, ethical, and military-operational concerns).

17. *Meaningful Human Control, Artificial Intelligence and Autonomous Weapons*, ARTICLE 36 (2016), <http://www.article36.org/wp-content/uploads/2016/04/MHC-AI-and-AWS-FINAL.pdf> [hereinafter Article 36, MHC] (“At its most basic level, the requirement for MHC develops from two premises: (1) That a machine applying force and operating without any human control whatsoever is broadly considered unacceptable. (2) That a human simply pressing a ‘fire’ button in response to indications from a computer, without cognitive clarity or awareness, is not sufficient to be considered ‘human control’ in a substantive sense.”). *cf.* Jeffrey S. Thurnher, *No One at the Controls: Legal Implications of Fully Autonomous Targeting*, 67 JOINT FORCE Q., 4th Quarter at 77, 83 (2012) (noting that oversight would not be effective if the human operator were merely just a rubber stamp to approve an engagement); M.C. Elish, *Moral Crumple Zones: Cautionary Tales in Human-Robot Interaction* (Mar. 20, 2016) (unpublished manuscript) (on file with author) (discussing how individuals may become “moral crumple zones” in accidents involving human/algorithmic interactions).

18. See Eric Messinger, *Is It Possible to Ban Autonomous Weapons in Cyberwar?*, JUST SEC. (Jan. 15, 2015), <https://www.justsecurity.org/19119/ban-autonomous-weapons-cyberwar/> (“The nature of the battlefield in question and the character of cyber defense will dictate the capabilities necessary for effective offensive operations, and the trend will be toward autonomous systems.”).

19. *Id.*

20. See, e.g., *Key Elements of Meaningful Human Control*, ARTICLE 36 (2016), <http://www.article36.org/wp-content/uploads/2016/04/MHC-2016-FINAL.pdf> [hereinafter Article 36, Key Elements]; Article 36, MHC, *supra* note 17.

21. Frank Sauer, *ICRAC Statement on Technical Issues to the 2014 UN CCW Expert Meeting*, ICRAC INT’L COMM. FOR ROBOT ARMS CONTROL (May 14, 2014), [icrac.net/2014/05/icrac-statement-on-technical-issues-to-the-un-ccw-expert-meeting](http://icrac.net/2014/05/icrac-statement-on-technical-issues-to-the-un-ccw-expert-meeting) [hereinafter ICRAC Statement].

1. “[A] human commander (or operator) must have full contextual and situational awareness of the target area and be able to perceive and react to any change or unanticipated situations that may have arisen since planning the attack”;
2. “[T]here must be active cognitive participation in the attack and sufficient time for deliberation on the nature of the target, its significance in terms of the necessity and appropriateness of attack, and likely incidental and possible accidental effects of the attack”; and
3. “[T]here must be a means for the rapid suspension or abortion of the attack.”<sup>22</sup>

Soon after, the Center for a New American Security (CNAS) reviewed the existing literature and concluded that there are three “essential components” to meaningful human control:

1. Human operators are making informed, conscious decisions about the use of weapons.
2. Human operators have sufficient information to ensure the lawfulness of the action they are taking, given what they know about the target, the weapon, and the context for action.
3. The weapon is designed and tested, and human operators are properly trained, to ensure effective control over the use of the weapon.<sup>23</sup>

More recently, Article 36 has argued that “meaningful human control” requires

1. “[p]redictable, reliable and transparent technology”;
2. “[a]ccurate information for the user on the outcome sought, the technology, and the context of use”;
3. “[t]imely human judgement and action, and a potential for timely intervention”; and
4. “[a]ccountability to a certain standard.”<sup>24</sup>

The International Committee of the Red Cross, meanwhile, has stated that “meaningful human control” entails “strict operational constraints with respect to the task carried out, the targets attacked, the operational environment, the geographical space and time of operation, the scope to enable human oversight of the weapon system, and the human ability to deactivate it if need be.”<sup>25</sup>

Drilling down on the relevant factors shifts the terms of the debate in a productive way—but it does not eliminate the principle’s inherent imprecision. With regard to the CNAS definition, for example, instead of arguing over what “meaningful” or “control” demands, States will likely debate what constitutes “informed decisions,” “sufficient information,” or “proper training.”<sup>26</sup>

22. *Id.*

23. Horowitz & Scharre, *supra* note 5, at 4.

24. Article 36, Key Elements, *supra* note 20, at 1.

25. ICRC Statement, *supra* note 13.

26. Additionally, as Marc Canellas and Rachel Haga point out, significantly more precision will be necessary before the standard can be practically applied at the design level. Marc C.

Additionally, broader questions remain unresolved. Should control be exercised over the entire weapon system, or perhaps only over its “critical functions”?<sup>27</sup> If so, what functions are “critical”?<sup>28</sup> Alternatively, perhaps meaningful human control must be exercised over individual attacks, regardless of whether an AWS may be employed? If so, there are different aspects of an attack, each of which may be subject to different degrees of human control, including (1) why someone or something is targeted; (2) how force is used; (3) who or what is harmed by the attack (both directly and indirectly); (4) when force is applied or harm is experienced; and (5) where force is applied and harm is experienced.<sup>29</sup> Furthermore, there is still no agreement as to the level of decision-making at which human control must occur.<sup>30</sup> The commander determining the rules of engagement is exercising a certain kind of control, the commander ordering a particular attack is exercising another, and the individual implementing that order might exercise yet another kind of control.<sup>31</sup>

Given the difficulty in pinning down what “meaningful human control” actually requires, “[s]everal states [have] expressed skepticism over the added value of the suggested concept, assessing it as being too vague, subjective and unclear.”<sup>32</sup>

### III. BENEFITS TO IMPRECISION IN INTERNATIONAL TECHNOLOGY REGULATION

It may well be, however, that the principle’s indefiniteness is its strength. International law is built on State consensus, and it is often easier to get States to first agree to a progressive but vague statement or principle—say, that everyone has the right to life<sup>33</sup>—and later hash out what it actually entails, given different stances on issues like abortion and capital punishment.<sup>34</sup> Meaningful human

Canellas & Rachel A. Haga, *Toward Meaningful Human Control of Autonomous Weapons Systems Through Function Allocation*, INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS, 6 (2015) (suggesting that “the effective function allocation literature provides a framework for designing rules and standards for AWS”).

27. See UNIDIR REPORT, *supra* note 3, at 2 (questioning whether meaningful human control should be exercised over weapons systems, over critical functions of autonomous weapons, or over individual attacks).

28. See ICRC REPORT, *supra* note 1, at 44 (defining “critical functions” as those relating to “selecting and attacking targets”).

29. See Horowitz & Scharre, *supra* note 5, at 15 (raising similar questions); see also Maya Brehm, Presentation to the Informal Meeting of Experts on Lethal Autonomous Weapon Systems 1–2 (Apr. 14, 2015), [http://www.unog.ch/80256EDD006B8954/\(httpAssets\)/897D1C5358C307BDC1257E280028024B/\\$file/BREHM\\_Presentation+on+MHC\\_14.04.2015.pdf](http://www.unog.ch/80256EDD006B8954/(httpAssets)/897D1C5358C307BDC1257E280028024B/$file/BREHM_Presentation+on+MHC_14.04.2015.pdf) (highlighting different aspects of an attack over which human control may be exercised).

30. Horowitz & Scharre, *supra* note 5, at 15.

31. *Id.*

32. Biontino Report, *supra* note 2, at 13; see also *id.* at 17 (“Meaningful human control may be useful as a policy approach to address shortcomings in current technology. However, it should not be applied as a legal criterion as this could undermine existing targeting law by introducing ambiguity.”).

33. Universal Declaration of Human Rights art. 3, G.A. Res. 217 (III) A, U.N. Doc. A/RES/217(III) (Dec. 10, 1948).

34. Cf. Brian Israel, *Treaty Stasis*, AJIL UNBOUND (May 8, 2014), <http://asil.org/blogs/treaty-stasis-agera-end-treaties> (“Relatively open-textured treaties that

control, as a phrase, is particularly useful in that it invites commentary and interpretation from a wide variety of stakeholders, including State representatives, weapon designers and manufacturers, human rights activists, philosophers, and (of course) lawyers.<sup>35</sup> Additionally, flexible terms that simultaneously draw a line prohibiting certain extreme developments while allowing for adaptive interpretations are of particular use in law intended to regulate new technology, especially weapons technologies.<sup>36</sup>

Take, for example, the two existing and similarly imprecise restrictions on new weaponry.<sup>37</sup> A weapon cannot be “by nature indiscriminate,” which is to say, it must be capable of being used in a way that discriminates between lawful targets (combatants, civilians directly participating in an attack, and other military objectives) and unlawful targets (civilians, civilian objects, and incapacitated or surrendering combatants).<sup>38</sup> Nor can a weapon “cause superfluous injury or unnecessary suffering.”<sup>39</sup> Both of these requirements are appealing in the abstract

---

prescribegeneral principles supply a framework for answering the governance questions that inevitably arise over a treaty’s lifetime with the advent of capabilities and activities not expressly addressed by the treaty.”).

35. Cf. Article 36, Key Elements, *supra* note 20, at 2 (noting that any modifier for “human control” would be ambiguous, but that “meaningful” is preferable because “it is general rather than context specific (e.g. appropriate), derives from an overarching principle rather [than] being outcome driven (e.g. effective, sufficient), and it implies human meaning rather than something administrative, technical or bureaucratic”); UNIDIR REPORT, *supra* note 3, at 3 (“[‘Meaningful human control’] provides a common language for discussion that is accessible to a broad range of governments and publics regardless of their degree of technical knowledge. . . . It focuses on a potentially shared objective of maintaining some form of control over all weapon systems . . . . It is a concept broad enough to integrate consideration of ethics, human-machine interaction and the ‘dictates of the public conscience’ which are often side-lined in approaches that narrowly consider technology, law or functions.”).

36. Cf. Colin B. Picker, *A View from 40,000 Feet: International Law and the Invisible Hand of Technology*, 23 CARDOZO L. REV. 149, 185–86 (2001); Charles J. Dunlap, *To Ban New Weapons or Regulate Their Use?*, JUST SECURITY (Apr. 3, 2015), <https://www.justsecurity.org/21766/guest-post-ban-weapons-regulate-use/> (arguing for regulating weapons’ effects, rather than weapons themselves). *But see* Chris Reed, *Taking Sides on Technology Neutrality*, 4 SCRIPT-ED 264, 280 (2007) (questioning the achievability and desirability of tech-neutral regulations).

37. States Parties to the Geneva Conventions developing or acquiring new weapons are required “to determine whether its employment would, in some or all circumstances, be prohibited by this Protocol or by any other rule of international law applicable.” Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of International Armed Conflicts (Protocol I) art. 36, *adopted* June 8, 1977, 1125 U.N.T.S. 3 [hereinafter First Additional Protocol]. Many argue that this responsibility is one of customary law, as it “flows logically from the truism that States are prohibited from using illegal weapons, means and methods of warfare or from using weapons, means and methods of warfare in an illegal manner.” Int’l Comm. of the Red Cross, *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).

38. *See, e.g.*, First Additional Protocol, *supra* note 37, art. 51(4); *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](https://www.icrc.org/customary-ihl/eng/docs/v1_rul_rule71) (last visited Aug. 2, 2015).

39. *See, e.g.*, Declaration Renouncing the Use, in Time of War, of Certain Explosive



and clearly prohibit or permit the use of certain weapons at the extremes, but they also allow for a fair amount of good faith disagreement about the lawfulness of a weapon in the middle ground. For example, when advocates for a ban on anti-personnel landmines argued that they were already forbidden as inherently indiscriminate,<sup>40</sup> the United States countered that “smart” anti-personnel mines, which self-destruct or deactivate automatically, are not *per se* unlawful.<sup>41</sup>

Notwithstanding—or perhaps because of—its inherent imprecision, a proscription on weapons incapable of being meaningfully controlled might usefully join existing prohibitions on new weaponry. While incorporating this principle in States’ legal reviews will not address all of the issues raised by AWS,<sup>42</sup> it will help concretize what meaningful human control actually means through State practice. However, this principle will only be beneficial if it does not undermine existing humanitarian norms governing targeting.

#### IV. AN INTERPRETATIVE FLOOR

All lawful attacks in an armed conflict must distinguish between lawful and unlawful targets (the distinction requirement);<sup>43</sup> must not cause excess collateral damage relative to the anticipated military advantage (the proportionality

Projectiles, *adopted* Dec. 11, 1868, 18 Martens Nouveau Recueil (ser. 1) 474, 138 Consol. T.S. 297; First Additional Protocol, *supra* note 37, art. 35(2); *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](http://www.icrc.org/customary-ihl/eng/docs/v1_rul_rule70) (last visited Aug. 2, 2015).

40. *See, e.g., Why the Ban*, INT’L CAMPAIGN TO BAN LANDMINES, <http://www.icbl.org/eng/problem/why-the-ban.aspx> (last visited Aug. 2, 2015).

41. *See* Emily Alpert, *Why Hasn’t the U.S. Signed an International Ban on Landmines?*, L.A. TIMES BLOG (Apr. 5, 2012), [http://latimesblogs.latimes.com/world\\_now/2012/04/mine-treaty-us-ottawa-convention.html](http://latimesblogs.latimes.com/world_now/2012/04/mine-treaty-us-ottawa-convention.html) (discussing allusions by U.S. officials that the use of smart land mines eliminated potential danger to civilians and therefore had a legitimate battlefield use).

42. Some have suggested that the meaningful human control principle could solve the AWS accountability problem. *Cf. Killer Robots and the Concept of Meaningful Human Control*, HUM. RTS. WATCH & INT’L HUMAN RIGHTS CLINIC, HARVARD LAW SCH. 6 (2016), [https://www.hrw.org/sites/default/files/supporting\\_resources/robots\\_meaningful\\_human\\_control\\_final.pdf](https://www.hrw.org/sites/default/files/supporting_resources/robots_meaningful_human_control_final.pdf) (“An obligation to have meaningful human control would allow for the imposition of legal liability and avoid the accountability gap associated with fully autonomous weapons.”).

It will not, unless the current international criminal law standard is changed. A human operator might make an informed decision, based on sufficient information and training, to deploy a well-tested and largely predictable AWS—and, due to its capacity for independent and thus sometimes unpredictable action, the AWS might nonetheless act in a way that results in a serious violation of international humanitarian law without anyone involved acting willfully. At that point, it will be possible to identify which person or people were in putative control of the system, but as no one—not the deployer, operator, commander, programmer, developer, manufacturer, or the weapon system itself—can be held liable under international criminal law absent willful action, the accountability gap remains.

The AWS accountability gap would be best addressed by constructing an international liability regime of “war torts” to deter serious violations of international humanitarian law, improve formal mechanisms for the recognition of state fault, and provide compensation to victims. Rebecca Crootof, *War Torts: Accountability for Autonomous Weapons*, 164 U. PA. L. REV. 1347 (2016).

43. First Additional Protocol, *supra* note 37, art. 48.

requirement);<sup>44</sup> and must minimize collateral damage and incidental injury (the feasible precautions or humanity requirement).<sup>45</sup> In many situations, meaningful human control will usefully augment these humanitarian norms.<sup>46</sup> The more control an operator exercises over a weapon system or attack, for example, the more likely she is able to take feasible precautions. This principle is potentially destabilizing, however, to the extent it might be interpreted to conflict with existing norms and suggest that they could be sacrificed in the name of ensuring additional control.

Consider the ICRC's definition of meaningful human control, which requires the human commander or operator to have "full contextual and situational awareness of the target area" and the "means for the rapid suspension or abortion of the attack."<sup>47</sup> If these "minimum necessary" requirements applied to all attacks, many weapons that have historically or are currently being employed would be rendered unlawful, to the detriment of both soldiers and civilians. As CNAS notes, "humans have been employing weapons where they lack perfect, real-time situational awareness of the target area since at least the invention of the catapult" and "the essence of a projectile weapon, since the first time a human hurled a rock in anger, is the inability to suspend and abort the attack after launch."<sup>48</sup> Additionally, defensive AWS are uniquely effective precisely because they can identify, track, target, and engage an incoming threat, possibly before a human being even knows of its existence.<sup>49</sup> In such situations, requiring "active cognitive participation" might expose troops to unnecessary levels of risk.<sup>50</sup>

Not only does the ICRC's definition "articulate an idealized version of human control divorced from the reality of warfare,"<sup>51</sup> it actually threatens to undermine fundamental humanitarian norms governing targeting.<sup>52</sup> Consider a commander deciding between employing a precision-guided munition and a

---

44. *Id.* art. 51(5)(b).

45. *Id.* art. 57(2)(a)(ii).

46. *Cf.* UNIDIR REPORT, *supra* note 3, at 4 (observing that the principle of meaningful human control usefully "turns our attention away from speculation about technological development and future capabilities and toward articulating the expectation that the development and use of emerging technologies will conform to established norms of responsibility, accountability, legality, and other principles of international humanitarian and human rights law").

47. ICRC Statement, *supra* note 21.

48. Horowitz & Scharre, *supra* note 5, at 9.

49. *Cf.* Crootof, *supra* note 1, at 1849 (discussing defensive autonomous weapon systems in use today).

50. I draw on this scenario to argue that certain weapon systems are "effectively autonomous" when a human operator does not have "sufficient time to evaluate the nature of the target, its military significance, and the likely incidental effects of engagement." *Id.* at 1858–61.

51. Horowitz & Scharre, *supra* note 5, at 9.

52. Kenneth Anderson & Matthew C. Waxman, *Debating Autonomous Weapon Systems, Their Ethics and Their Regulation Under International Law*, in OXFORD TECHNOLOGY HANDBOOK (forthcoming 2017) (manuscript on file with author), at 23 ("[A]lthough some of its proponents view the MHC standard as flowing from LOAC, in some important respects it is quite at odds with the fundamental structure of LOAC, and its core principles of necessity, distinction, proportionality, and humanity.").

human-piloted bomber to attack a target. In many cases, the former would be far preferable for both sides in a conflict. Not only does increased distance between troops and a target reduce physical and psychological risk to the side orchestrating the attack, the accuracy of precision-guided munitions reduces the need to use excessive lethal force when attacking from afar, which in turn lessens the likelihood of collateral damage.<sup>53</sup> Similarly, “autonomous weapon systems promise a next-generation combination of distance, accuracy, and lethality”—which may decrease risks both to the combatants fielding an AWS and to civilians.<sup>54</sup> These potential gains in human safety should not be sacrificed to an overly-strict interpretation of what constitutes meaningful human control.

## V. CONCLUSION

In large part because of its inherent imprecision, the principle of “meaningful human control” can fruitfully advance the conversation regarding the appropriate regulation of autonomous (and other) weapon systems, especially if it augments existing humanitarian norms governing targeting. But these norms—the distinction, proportionality, and feasible precaution requirements—should serve as an interpretative floor.<sup>55</sup> Any definition of meaningful human control that would prioritize human control at the cost of increased risk to soldiers and civilians must be rejected outright.

---

53. See Oren Gross, *The New Way of War: Is There a Duty to Use Drones?*, 67 FLA. L. REV. 1, 30, 47–48 (2015) (discussing the relationship between distance and necessary amount of lethal force).

54. Crotoof, *War Torts*, *supra* note 42, at 1370-71.

55. Kenneth Anderson, Daniel Reisner & Matthew Waxman, *Adapting the Law of Armed Conflict to Autonomous Weapon Systems*, 90 INT'L L. STUD. 386, 410 (2014) (“The fundamental principle underlying the gradual development of [international humanitarian law] standards and rules alongside the evolution of automation technologies, however, should be that what matters is ever greater compliance with the core obligations of the law of armed conflict: necessity, distinction, proportionality, and humanity.”).