Bombs, Bots, and the Principle of Distinction: The Law of Armed Conflict and Contemporary Warfare

Nathan G. Wood

Texas National Security Review 9, no. 1 (Winter 2026) | © 2026 by Nathan G. Wood | DOI: 10.1353/tns.00023

Critics of autonomous weapons systems (AWS) argue that these weapons cannot reliably distinguish between legitimate targets and those protected from attack. As a result, the use of AWS seems to violate the principle of distinction under international humanitarian law (IHL), which requires that combatants "not make civilians the object of attack" and not carry out attacks that are "indiscriminate in nature." This criticism, however, misunderstands the principle of distinction and ignores important aspects of how AWS are being developed and deployed. Critics rely on an overly broad definition of AWS, and hold these systems to a standard that is inconsistent with the principle of distinction as it is actually formulated under IHL. Despite their very real limitations, the characteristics of modern AWS in fact highlight an impressive feat of technological design, and mark a further step on our long and fitful road to making warfare a less brutal and bloody enterprise.

Militaries around the world are developing increasingly autonomous weapons systems. These efforts, however, have been met with staunch opposition from a number of groups.[1] Critics object that "autonomous weapons would face great, if not insurmountable, difficulties in reliably distinguishing between lawful and unlawful targets as required by international humanitarian law."[2] One part of this objection argues that autonomous weapons system(s) (AWS) cannot recognize the subtle cues that can distinguish civilians from combatants, active combatants from those attempting to surrender, or active combatants from those who are out of action due to unconsciousness, wounds, or sickness (hors de combat).[3] A second concern is that autonomous weapons may have certain features that render them unpredictable, making their use inherently indiscriminate.[4] As a result, opponents charge, such weapons would be in breach of international humanitarian law (IHL) and the principle of distinction.[5]

These objections fundamentally mistake what is required under the principle of distinction, rely on inaccurate depictions of AWS, and misunderstand what IHL demands with regard to the use of these weapons. The core of the principle of distinction is not concerned with regulating weapons, but rather uses of weapons. Precautions in attack and the principle of proportionality both affect distinction in war, and the relevant question is not about the technology per se, but the conditions under which commanders are required to exercise greater care.

1 Most notably, the Campaign to Stop Killer Robots, Human Rights Watch, and Amnesty International, as well as numerous scholars.

2 Human Rights Watch, *Making the Case: The Dangers of Killer Robots and the Need for a Preemptive Ban* (Human Rights Watch, International Human Rights Clinic at Harvard Law School, 2016), 5.

3 Noel Sharkey, "Saying 'No!' to Lethal Autonomous Targeting," *Journal of Military Ethics* 9, no. 4 (2010): 369–83, https://doi.org/10.1080/15027570.2010.537903; Marcello Guarini and Paul Bello, "Robotic Warfare: Some Challenges in Moving from Noncivilian to Civilian Theaters," in *Robot Ethics: The Ethics and Social Implications of Robotics*, ed. Patrick Lin, Keith Abney, and George A. Bekey (MIT Press, 2012); Human Rights Watch, *Losing Humanity: The Case Against Killer Robots* (Human Rights Watch, International Human Rights Clinic at Harvard Law School, 2012); Robert Sparrow, "Twenty Seconds to Comply: Autonomous Weapon Systems and the Recognition of Surrender," *International Law Studies* 91, no. 1 (2015): 699–728, https://digital-commons.usnwc.edu/ils/vol91/iss1/20/; Robert Sparrow, "Robots and Respect: Assessing the Case Against Autonomous Weapon Systems," *Ethics & International Affairs* 30, no. 1 (2016): 93–116, https://doi.org/10.1017/s0892679415000647; Elliot Winter, "The Compatibility of Autonomous Weapons with the Principle of Distinction in the Law of Armed Conflict," *International & Comparative Law Quarterly* 69, no. 4 (2020): 845–76, https://doi.org/10.1017/s0020589320000378; Stop Killer Robots, *Negotiating a Treaty on Autonomous Weapons Systems: The Way Forward* (Stop Killer Robots, 2022).

4 Alexander Blanchard and Mariarosaria Taddeo, "Predictability, Distinction & Due Care in the Use of Lethal Autonomous Weapon Systems," *SSRN Electronic Journal* (2022), https://dx.doi.org/10.2139/ssrn.4099394.

5 In some discussions, especially those outside of legal discourse, distinction is referred to under the heading of "discrimination." Some analysts, however, defend the view that distinction and discrimination refer to separate legal demands. See, for example, Adil Ahmad Haque, *Law and Morality at War* (Oxford University Press, 2017), 106.

The principle of distinction will indeed set bounds on how AWS may be used, and on when, where, and under what limitations they may be deployed, but it cannot underpin any blanket prohibitions for existing or near-future autonomous weapons.[6] Autonomous weapons do create legitimate ethical and legal worries, but we must identify actual problems rather than chasing phantom concerns rooted in misunderstandings about technology and military operations

Laying the Groundwork: Autonomous Weapons Systems and International Law

To begin, a clear definition of autonomous weapon systems is necessary to avoid misunderstanding.[7] The predominant view is that autonomous weapon systems are those that, once activated, are capable of selecting and engaging targets without contemporaneous human input. Variations of this view appear in definitions from the US Department of Defense, the International Committee of the Red Cross (ICRC), Human Rights Watch, and critics of AWS.[8] This definition is broad, including not just futuristic systems but other mundane platforms that have existed for decades, such as the close-in weapon systems used by most navies around the world for missile interception. The ongoing Russo-Ukrainian war has also seen broad use of a variety of increasingly autonomous systems for numerous roles, underpinning the fact that AWS are not just systems of tomorrow, but of yesterday and today.[9] Thus, under this expansive definition, not only would Terminators (from the science fiction films) count as AWS, but so would anti-radiation missiles, certain anti-armor weapons, autonomous point-defense systems, and many forms of loitering or self-guided munitions.[10]

Critics' core objection to the use of AWS is that these weapons will not be able to reliably distinguish between lawful and unlawful targets because technical limitations in computer systems make them unable to exercise necessary human-level judgment.

6 An early argument similar to this can be found in Michael N. Schmitt, "Autonomous Weapon Systems and International Humanitarian Law: A Reply to the Critics," *Harvard National Security Journal Feature* (2012), https://dx.doi.org/10.2139/ssrn.2184826, which is in response to Human Rights Watch, *Losing Humanity*, *supra* note 3. In the aftermath of that exchange, voices critical of autonomous weapons were careful for a time to distinguish between *systems* and *deployments of systems* (see, for example, Christof Heyns, *A/HRC/23/47 Report of the Special Rapporteur on Extrajudicial, Summary or Arbitrary Executions* [United Nations General Assembly, 2013], https://www.ohchr.org/sites/default/files/Documents/HRBodies/HRCouncil/RegularSession/Session23/A-HRC-23-47_en.pdf), but within only a few years (see, for example, Sparrow, "Twenty Seconds to Comply," *supra* note 3; or Human Rights Watch, *Making the Case*, *supra* note 2), the conflation of concerns had again arisen, and the contemporary debate often displays deep confusion on this point.

7 Nathan G. Wood, "Autonomous Weapon Systems: A Clarification," *Journal of Military Ethics* 22, no. 1 (2023): 18–32, https://doi.org/10.1080/15027570.2023.2214402. For a schematic approach, see Andrew P. Williams, "Defining Autonomy in Systems: Challenges and Solutions," in *Autonomous Systems: Issues for Defense Policymakers*, ed. Andrew P. Williams and Paul Scharre (NATO Communications and Information Agency, 2015), 27. For examinations of various definitions and their components, see also William H. Boothby, *Weapons and the Law of Armed Conflict*, 2nd ed. (Oxford University Press, 2016); Jürgen Altmann and Frank Sauer, "Autonomous Weapon Systems and Strategic Stability," *Survival* 59, no. 5 (2017), https:/doi.org/10.1080/00396338.2017.1375263; Jean-François Caron, "Defining Semi-Autonomous, Automated and Autonomous Weapon Systems in Order to Understand Their Ethical Challenges," *Digital War* 1, nos. 1–3 (2020), https://doi.org/10.1057/s42984-020-00028-5; Mariarosaria Taddeo and Alexander Blanchard, "A Comparative Analysis of the Definitions of Autonomous Weapons Systems," *Science and Engineering Ethics* 28, no. 5 (2022): 1–22, https:/doi.org/10.1007/s11948-022-00392-3; Magdalena Pacholska, "Autonomous Weapons," in *Research Handbook on Law and Technology*, ed. Bartosz Brożek, Olia Kanevskaia, and Przemysław Pałka (Edward Elgar, 2024). For the state of the ongoing definitional debates being had within the Group of Governmental Experts (GGE) working under the auspices of the Convention on Conventional Weapons, see https://meetings.unoda.org/ccw/convention-on-certain-conventional-weapons-group-of-governmental-experts-on-lethal-autonomous-weapons-systems-2025. That definition also includes significant overlap with the definition given here, with the caveat that the GGE's definition is of "LAWS," or *lethal* autonomous weapon systems, and as such, they place an explicit focus on lethality of systems, even systems being used for anti-materiel purposes.

8 United States Department of Defense, *DoD Directive 3000.09: Autonomy in Weapon Systems* (United States Department of Defense, 2023), 21; International Committee of the Red Cross, *Autonomous Weapons Systems: Technical, Military, Legal and Humanitarian Aspects* (International Committee of the Red Cross, 2014), 5; International Committee of the Red Cross, *ICRC Position on Autonomous Weapons Systems* (International Committee of the Red Cross, 2021), 1; Human Rights Watch, *New Weapons, Proven Precedent: Elements of and Modes for a Treaty on Killer Robots* (Human Rights Watch, International Human Rights Clinic at Harvard Law School, 2020), 2; Peter Asaro, "On Banning Autonomous Weapon Systems: Human Rights, Automation, and the Dehumanization of Lethal Decision-Making," *International Review of the Red Cross* 94, no. 886 (2012): 690, https://doi.org/10.1017/S1816383112000768; Heather M. Roff, "The Strategic Robot Problem: Lethal Autonomous Weapons in War," *Journal of Military Ethics* 13, no. 3 (2014): 211–27, https://doi.org/10.1080/15027570.2014.975010; Thompson Chengeta, "Measuring Autonomous Weapon Systems Against International Humanitarian Law Rules," *Journal of Law & Cyber Warfare* 5 (2016): 67, https://dx.doi.org/10.2139/ssrn.2755184; Sparrow, "Robots and Respect," *supra* note 3, at 95; Peter Asaro, "Algorithms of Violence: Critical Social Perspectives on Autonomous Weapons," *Social Research: An International Quarterly* 86, no. 2 (2019): 540, https://doi.org/10.1353/sor.2019.0026; Stop Killer Robots, *Negotiating a Treaty on Autonomous Weapons Systems*, *supra* note 3, at 6.

9 See, for example, Vera Bergengruen, "How Tech Giants Turned Ukraine into an AI War Lab," *TIME*, February 8, 2024, https://time.com/6691662/ai-ukraine-war-palantir/; Vitaliy Goncharuk, *Survival of the Smartest? Defense AI in Ukraine* (Defense AI Observatory, Helmut Schmidt University, 2024); Jean-Marc Rickli and Federico Mantellassi, *The War in Ukraine: Reality Check for Emerging Technologies and the Future of War* (Geneva Centre for Security Policy, 2024); Michael C. Horowitz, *Ukraine's Operation Spider Web Shows Future of Drone Warfare* (Council on Foreign Relations, 2025), https://www.cfr.org/expert-brief/ukraines-operation-spider-web-shows-future-drone-warfare.

10 Vincent Boulanin, Neil Davison, Netta Goussac, and Moa Peldán Carlsson, *Limits on Autonomy in Weapon Systems: Identifying Practical Elements of Human Control* (Stockholm International Peace Research Institute [SIPRI] and International Committee of the Red Cross [ICRC], 2020); International Committee of the Red Cross, *ICRC Position and Background Paper on Autonomous Weapons Systems* (International Committee of the Red Cross, 2021); Nathan G. Wood, "Autonomous Weapon Systems and Responsibility Gaps: A Taxonomy," *Ethics and Information Technology* 25, no. 1 (2023): 1–14, https://doi.org/10.1007/s10676-023-09690-1; Wood, "Autonomous Weapon Systems," *supra* note 7.

In addition, critics argue that some AWS are too unpredictable for their general use to comply with the precautionary demands of distinction. Fundamentally, this objection is aimed at the weapons themselves. The problem, say critics, is not simply poor operational choices or flawed military judgment, but that the fundamental technical character of AWS means that they cannot satisfy distinction, making their use indiscriminate and unlawful.[11]

Distinction in IHL

To assess the objection fully, a clear definition of distinction is critical. A core principle in IHL, distinction sets basic requirements for legal conduct in war.[12] Aspects of distinction can be found in numerous treaties and throughout the Additional Protocols to the Geneva Conventions; in the ICRC's study of customary IHL, nearly a third of the rules involve the concept.[13] Distinction is often presented as a principle protecting civilians from attack and the effects of attack,[14] but at a basic level, it demands that combatants distinguish between lawful and unlawful targets and direct their attacks only at the former. This language obviously implies protection for civilians, but it also incorporates protections for combatants who are surrendering or otherwise out of combat.

Civilians

The "basic rule" of distinction with regard to civilians is embodied in Article 48 of Geneva Protocol I Additional to the Geneva Conventions (AP I), which states:

> *In order to ensure respect for and protection of the civilian population and civilian objects, the Parties to the conflict shall at all times distinguish between the civilian population and combatants and between civilian objects and military objectives and accordingly shall direct their operations only against military objectives.*

This rule is, according to the ICRC Commentary on the *Additional Protocols*, "the foundation on which the codification of the laws and customs of war rests,"[15] and is given substance and force throughout Part IV of AP I.

Articles 51 and 57 are particularly relevant to the use of autonomous weapons in war. Article 51 makes clear that "the civilian population and individual civilians shall enjoy general protection against dangers arising from military operations" (AP I, Art. 51.1), and that they "shall not be the object of attack" (AP I, Art. 51.2). Most importantly for our purposes, the rule prohibits indiscriminate attacks, defined as:

> a. those which are not directed at a specific military objective;
> b. those which employ a method or means of combat which cannot be directed at a specific military objective; or
> c. those which employ a method or means of combat the effects of which cannot be limited as required by this Protocol;
> and consequently, in each such case, are of a nature to strike military objectives and civilians or civilian objects without distinction. (AP I, Art. 51.4)[16]

The definition of indiscriminate attacks moreover includes a proportionality condition. The definition treats as indiscriminate any "attack which 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" (AP I, Art. 51.5.a).

Article 57 adds precautionary standards that must be met for an attack to be lawful. It stipulates that "constant care shall be taken to spare the civilian population, civilians and civilian objects" (AP I, Art. 57.1). This addition, in turn, requires that combatants "do everything feasible to verify that the objectives to be attacked are neither civilians nor civilian objects"; choose their "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 refrain from launching attacks "which 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" (AP I, Art. 57.2).

11 See International Committee of the Red Cross, *Commentary on the Additional Protocols to the Geneva Conventions*, ed. Yves Sandoz, Christophe Swinarski, and Bruno Zimmerman (International Committee of the Red Cross, 1987), Art. 36, for the relevant legal standard concerning weapons review.

12 Gary D. Solis, *The Law of Armed Conflict: International Humanitarian Law in War*, 2nd ed. (Cambridge University Press, 2016), 309.

13 See https://ihl-databases.icrc.org/en/customary-ihl/v1.

14 For example, Haque, *Law and Morality at War*, *supra* note 5, chapter 5; Winter, "The Compatibility of Autonomous Weapons with the Principle of Distinction in the Law of Armed Conflict," *supra* note 3.

15 International Committee of the Red Cross, *Commentary on the Additional Protocols to the Geneva Conventions*, 598, paragraph 1863.

16 See also Boothby, *Weapons and the Law of Armed Conflict*, *supra* note 7, esp. chapter 6.

Importantly, none of these legal requirements rule out attacks that are expected to harm civilians, nor do any other rules of international law. Rather, civilians are protected from being directly targeted in attacks, attackers are required to take precautions minimizing expected harm to civilians, and attacks expected to cause disproportionate harm to civilians are prohibited. Additionally, combatants are prohibited from using methods or means that cannot be used in a discriminate manner. These requirements, however, still permit combatants to carry out attacks that may harm civilians, so long as the above restrictions are satisfied. Combatants are indeed permitted to carry out such attacks *in full knowledge that civilians will be harmed*, so long as the target being attacked is a military one, civilian harm is minimized to the extent possible, and civilian harm is proportionate to the military advantage gained.

Persons Hors de Combat

The principle of distinction applies to enemy personnel as well as civilians. In particular, IHL prohibits attacks on those who should be recognized to be hors de combat, which applies to an individual when:

a. he is in the power of an adverse Party;
b. he clearly expresses an intention to surrender; or
1. he has been rendered unconscious or is otherwise incapacitated by wounds or sickness, and therefore is incapable of defending himself; provided that in any of these cases he abstains from any hostile act and does not attempt to escape. (AP I, Art. 41.2)[17]

Distinction thus extends protections to those who are regularly combatants, but are temporarily noncombatants either by virtue of their own actions (surrendering persons), by their inability to fight (incapacitated persons), or because they have been taken into custody after surrender or retrieval by enemy forces (detained persons).[18] Importantly, Article 41 also does not merely protect those who are seen to hold hors de combat status, but rather extends to those who "should be recognized to be *hors de combat*" (AP I, Art. 41.1). It is thus irrelevant whether a given combatant *believes* some enemy individual is hors de combat—the question is whether the enemy "should be recognized" to hold that status. To answer that question, combatants (or combat systems) must be able to recognize and respond to expressed intentions to surrender. Those carrying out attacks must also either be able to recognize when an enemy is incapacitated, or be limited in such a way that they cannot carry out attacks unless an enemy is clearly not incapacitated.[19]

The Demands of Distinction in Practice

At a basic level, following distinction would seem to be straightforward. Combatants usually wear uniforms and carry arms openly, and military equipment and vehicles are often unmistakable. In contemporary conflict, however, a number of factors complicate the assessment: fighters may not be wearing distinctive dress, civilian vehicles or buildings may share similarities with military variants,[20] and some noncombatants protected from attack include members of the armed forces (for example, chaplains, medical personnel, and those hors de combat).[21] How then does distinction work in practice? And what implications does that process have for autonomous weapons systems?

To begin with, it is helpful to separate the discussion between international armed conflicts (IAC) and non-international armed conflicts (NIAC). The main

17 See also Hague Convention IV, esp. Art. 23, Art. 3 of Geneva Conventions I–IV, and the broader discussion of Rule 47 of the ICRC Customary IHL Database (https://ihl-databases.icrc.org/en/customary-ihl/v1/rule47).

18 David H. Lee et al., *Law of Armed Conflict Deskbook*, 5th ed. (International and Operational Law Department, United States Army Judge Advocate General's Legal Center and School, 2015), 137–38; David H. Lee et al., *Operational Law Handbook* (International and Operational Law Department, United States Army Judge Advocate General's Legal Center and School, 2015), 17–19.

19 Under some views, further attention must also be paid to condition (a), as being "in the power" of an adverse party is argued by some to not be limited only to cases where persons have been detained or apprehended, but instead takes the wider view that a "defenceless adversary is *hors de combat* whether or not he has laid down arms" (International Committee of the Red Cross, *Commentary on the Additional Protocols to the Geneva Conventions*, *supra* note 15, at 484, paragraph 1612). This position, however, is arguably incorrect, as no enemy is truly defenseless unless they are either already in one's power (that is, detained) or are incapacitated. Put simply, an "unarmed" individual is still very much capable of acting lethally, and, contrary to the *Commentary on the Additional Protocols*, no amount of "overwhelmingly superior firepower . . . can force the adversary to cease combat" (International Committee of the Red Cross, *Commentary on the Additional Protocols to the Geneva Conventions*, *supra* note 15, at 484, paragraph 1612). So long as they are not incapacitated and have not been detained, it is the adversary's choice whether or not they will continue combat, even when continuation can only result in their certain death. Thanks to an anonymous reviewer for pressing me to clarify this point.

20 Some civilian vehicles and sites may also explicitly be used by armed forces, even in regular conflicts. For example, due to equipment shortages, both Russia and Ukraine have used civilian vehicles for logistics, medical evacuation, and rapid assault forces. Parties to a conflict using civilian vehicles for such purposes on a regular basis are under an obligation to mark the vehicles accordingly, in order to prevent blurring the lines between combatants and civilians (AP I, Art. 44.3), but even when they do so, the mere use of regularly civilian vehicles can complicate target identification for both humans and machines.

21 Leslie C. Green, *The Contemporary Law of Armed Conflict*, 3rd ed. (Manchester University Press, 2017), 111.

relevant difference between these forms of warfare with regard to distinction and AWS is the extent to which individuals engaged in hostilities and military equipment are clearly recognizable and distinct from civilian individuals and objects.[22] In what follows, IAC will be understood as armed conflict in which the armed forces of belligerent parties wear a distinctive combat dress, fly the colors of their party, and otherwise demonstrate in a clear and obvious fashion that they are combatants regularly engaged in warfighting purposes. By contrast, NIAC include conflicts where one or more parties have fighters who are not easily distinguishable from the civilian population; they may be civilians who take up arms at particular moments, or fighters who intentionally eschew recognizable emblems. In other instances, there may not be clear combat dress (for example, combatants forced to wear civilian clothes because there are no uniforms to be had). These factors do not constitute the whole of what distinguishes IAC from NIAC, nor do they map cleanly to all instances of these types of conflict—as some IAC involve non-uniformed combatants and some NIAC are between clearly recognizable combatant groups[23]—but these categories capture core differences that are nonetheless relevant for autonomous weapons' compliance with the principle of distinction.[24]

Distinction in International Armed Conflict

In international armed conflicts (IAC) between peer adversaries, attritional fighting directed at materiel often plays a central role; this type of operation, in turn, impacts how combatants satisfy the distinction principle. When the aim is to destroy vehicles or equipment, the primary element of distinction can often be satisfied through a simple object recognition exercise:[25] Is this a tank, jet fighter, ammunition depot, bunker, or the like?[26] Most cases permit an attack on such an object. Distinction still demands that combatants ensure that potential risks to civilians in the vicinity are minimized, and that expected harms to civilians are not excessive. But the core of distinction—the act of *distinguishing* between legitimate and illegitimate targets—is a rather straightforward exercise in object recognition.[27]

IAC therefore open a wide space in which autonomous weapons can be used in a legal and ethical manner. Autonomous weapons in anti-materiel roles need only be able to reliably recognize the signatures of particular enemy platforms. Thus, anti-radiation missiles that lock onto radio signatures, anti-tank weapons that utilize high-frequency radar to locate and engage armored vehicles, anti-missile systems that take the speed and heading of aircraft as targeting parameters, or other similar systems are all predictable ways for AWS to select, track, and engage targets.[28] As long as these systems are reliable, their failure to hit the right targets (due to the presence of civilian objects matching the weapon's target profile) will also be fairly predictable.

For example, if certain anti-tank AWS are utilized in an area with civilian heavy-construction vehicles, some of these latter objects may be engaged by an autonomous weapon owing to how it makes its target selections.[29] Yet this outcome does not put the autonomous weapon itself in breach of distinction. Rather, it indicates that particular usages may be in breach, given the risks to civilians who may be expected to be in the AWS' area of operations. Whether the usage of that AWS leads to a breach is also not

22 For more thorough examinations of definitions of armed conflict and the practicalities of complying with the rules governing targeting, see Michael N. Schmitt, Yoram Dinstein, and Charles H. B. Garraway, *The Manual on the Law of Non-International Armed Conflict, with Commentary* (International Institute of Humanitarian Law, 2006); James Clancy and Chuck Crossett, "Measuring Effectiveness in Irregular Warfare," *The US Army War College Quarterly: Parameters* 37, no. 2 (2007), https://doi.org/10.55540/0031-1723.2351; Lawrence Freedman, "Defining War," in *The Oxford Handbook of War*, ed. Yves Boyer and Julian Lindley-French (Oxford University Press, 2012); William H. Boothby, *The Law of Targeting* (Oxford University Press, 2012).

23 As a stark example of the former scenario, one may consider World War II, in which civilians directly participated in hostilities on a large scale in some regions, despite this conflict having been an IAC. Some of the guerrilla fighting in Central America in the latter half of the twentieth century provides an example of the latter scenario. Thanks to an anonymous reviewer for pressing me to clarify these historical precedents.

24 Using somewhat looser language, we may also consider the differences between large-scale combat operations (LSCO) compared to counterinsurgency operations (COIN). In many cases, this distinction may ultimately prove more apposite. Given the complications brought forward by the laws of occupation that often attend COIN (but not necessarily NIAC), however, I will, for the sake of simplifying the discussion, stick to the IAC/NIAC distinction in what follows.

25 Critically, object recognition need not be limited to visual identification of military assets or platforms, but may make use of other forms of sensing as well, such as high-frequency radar, acoustic sensors, thermal imaging, and so forth. Multimodal sensing apparatuses make possible robust and reliable object recognition capabilities, which can greatly limit false targeting of civilian objects.

26 More precisely, AP I, Art. 52.2, specifies that "military objectives are limited to those objects which by their nature, location, purpose or use make an effective contribution to military action and whose total or partial destruction, capture or neutralization, in the circumstances ruling at the time, offers a definite military advantage."

27 Surrendering crews of vehicles and the wounded who may lie around them present additional challenges, but ones that are not insurmountable. See, for example, Maciek Zając, *Autonomous Weapon Systems from a Just War Theory Perspective* (PhD thesis, University of Warsaw, 2022), chapter 6.3.

28 Wood, "Autonomous Weapon Systems and Responsibility Gaps," *supra* note 10, at 4–6.

29 Some anti-tank systems scan for vehicles with particular acoustic or seismic signatures and these variables create possibilities of mistaken targeting of heavy construction or farming machinery in some situations.

a foregone conclusion. For example, destroying enemy armor in the region may be a militarily critical objective, the AWS may be the most discriminate weapon available for destroying that armor, and the risk to civilians, while real, may not outweigh the military advantage gained. In short, the demands of distinction will be contextual and nuanced, and the mere possibility of mistargeting civilian objects will not render the use of a particular AWS in breach of IHL.[30] Anti-materiel systems have limitations that make mistargeting possible in certain circumstances, but this quality does not undermine the overall legality of such weapons. Rather, these possible uses highlight the need for combatants to exercise good judgment in their choice of means and methods of attack, with a view to minimizing risks to civilians.

In addition to the destruction of equipment, warfare often necessitates the neutralization of enemy combatants.[31] In IAC, where combatants wear distinctive combat dress, the primary duty to distinguish between combatants and civilians is generally simplified.[32] This type of scenario means that autonomous weapon systems would need to be able to competently determine whether some article of clothing matches an enemy uniform, or at least be able to make a determination at a level comparable to that of a human combatant in a similar situation.[33] Machines already possess such capability,[34] so the primary demand of distinction can be satisfied by an AWS. Moreover, even if some particular AWS is less capable, whether its use violates distinction depends on whether or not *the attack for which it is used* is one where civilians are expected to be present, whether civilians are expected to be mistakenly targeted, whether civilian harms could not be mitigated through the use of another weapon, and whether civilian harms are excessive in comparison to the military advantage gained.

Two other categories that present potential problems for distinction are surrendering combatants and combatants rendered hors de combat due to sickness, wounds, or incapacitation. Critics argue that AWS will be incapable of recognizing or respecting combatants' attempts to surrender, thereby breaching distinction.[35] Many of these objections rely, either implicitly or explicitly, on the understanding that "differentiating between combatants and noncombatants will often require the appropriate attribution of mental states (such as intentions)," something that machines cannot (yet) do.[36] This assumption, however, is in error.

In short, the demands of distinction will be contextual and nuanced, and the mere possibility of mistargeting civilian objects will not render the use of a particular AWS in breach of IHL.

In the letter of international law, a combatant is not to be targeted when he or she "clearly expresses an intention to surrender" (AP I, Art. 41.2.b). The requirement is that a surrendering combatant needs to *clearly express* that intention; the opposing side does not have the responsibility to divine the secret intentions of their enemies, or to ascertain the mental states of those they are fighting. According to the ICRC's *Commentary on the Additional Protocols,* combatants wishing to surrender may carry out any

30 For a separate line of argumentation that reaches a similar conclusion, see Magdalena Pacholska, "Military Artificial Intelligence and the Principle of Distinction: A State Responsibility Perspective," *Israel Law Review* 56, no. 1 (2023): 3–23, https://doi.org/10.1017/s0021223722000188.

31 For novel recent resistance to this assumed position, see Ido Rosenzweig, "'When You Have to Shoot, Shoot!' Rethinking the Right to Life of Combatants During Armed Conflicts," *International Review of the Red Cross* 106, no. 926 (2024): 863–95, https://doi.org/10.1017/s181638312400016x. Discussion of this position is unfortunately beyond the scope of this article.

32 Note, though, that some IAC may involve large-scale civilian direct participation in hostilities, complicating matters. See note 23.

33 Kevin Jon Heller, "The Concept of 'The Human' in the Critique of Autonomous Weapons," *Harvard National Security Journal* 15, no. 1 (2023): 1–76, esp. 19–20, https://harvardnsj.org/wp-content/uploads/2023/12/Heller_15-Harv.-Natl-Sec.-J.-1-2023.pdf.

34 Paul Scharre and Kelley Sayler, *Autonomous Weapons and Human Control* (Center for a New American Security, 2016); Winter, "The Compatibility of Autonomous Weapons with the Principle of Distinction in the Law of Armed Conflict," *supra* note 3, at 867.

35 For example, Noel Sharkey, "Grounds for Discrimination: Autonomous Robot Weapons," *RUSI Defence Systems* 11, no. 2 (2008): 86–89; Noel Sharkey, "Killing Made Easy: From Joysticks to Politics," in *Robot Ethics: The Ethics and Social Implications of Robotics*, ed. Patrick Lin, Keith Abney, and George A. Bekey (MIT Press, 2012); Human Rights Watch, *Losing Humanity*, *supra* note 3; Chengeta, "Measuring Autonomous Weapon Systems Against International Humanitarian Law Rules," *supra* note 8. See also Eliav Lieblich and Eyal Benvenisti, "The Obligation to Exercise Discretion in Warfare: Why Autonomous Weapon Systems Are Unlawful," in *Autonomous Weapons Systems: Law, Ethics, Policy*, ed. Nihal Bhuta et al. (Cambridge University Press, 2014), 278–82, for a more nuanced and theoretical objection to AWS grounded in duties to respect surrender.

36 Guarini and Bello, "Robotic Warfare," *supra* note 3, at 129. Note that though Guarini and Bello's argument is utilized by many critics of AWS, their exact position is not that AWS cannot satisfy the demands of IHL, but rather that "until much more progress is made, we should not be sanguine about the advantages of robots in theaters on the *noncombatant end of the spectrum*" (emphasis in original).

number of internationally recognized actions: They may lay down arms and raise their hands, they may hoist a white flag, or they may verbally communicate their wish to surrender.[37] Even aircraft and ships at sea can signal surrender by, respectively, waggling their wings and lowering flags.[38] In any of these cases, it is the responsibility of surrendering combatants to make their intentions clear. Critically, recognized gestures of surrender constitute, or are accompanied by, deeds that actually strip the surrendering combatant of their ability to harm their enemies; that is, if one continues to fight, simultaneously raising a white flag or shouting "I surrender" does not automatically protect one from attack.[39] These patterns—such as the laying down of arms and raising of hands—can be reliably recognized by machines; in fact, some have already been integrated into existing autonomous platforms.[40]

A second potential problem of distinction in IAC involves soldiers who are hors de combat.[41] International law makes clear that persons who "*should* be recognized to be *hors de combat*" due to incapacitation shall not be made objects of attack. Recognizing whether one is incapacitated, however, can be difficult, especially in the fog of war and under rapidly changing circumstances, and will be highly contingent on who is making the assessment. For example, the infantrymen clearing a trenchwork have very different information than the helicopter crew providing close combat air support—differences in context and access to information that will impact legal assessments. Infantrymen clearing positions will likely be in a better position to determine whether or not enemies really are hors de combat, or should be recognized to be such; as a result, these soldiers will be required to not make certain wounded enemies the object of attack. Pilots, artillerymen, or other remote fighters will likely not possess such knowledge of the situation and will not be in a position to as reliably recognize wounded or incapacitated enemies as hors de combat—*nor will they be required to*. Put bluntly, just because a certain method or means of attack makes it difficult to determine an enemy's status does not ipso facto render that method or means illegal.

One may object that those who seem to be clearly incapacitated—for example, bleeding profusely, screaming in pain, holding body parts rather than weapons—obviously should be recognized to be hors de combat, regardless of who might be engaging them. An enemy missing a limb and lacking weapons can, arguably, hardly be anything but hors de combat. This objection, however, still ignores a critical point on the realities of war, and misses an important aspect of how actions in war are judged and what sorts of evidence are taken into consideration when applying the principle of distinction.

The central element in IHL is not that an individual is wounded or otherwise harmed, but rather whether an individual "therefore is incapable of defending himself."[42] A wounded enemy who continues to fight remains targetable until he or she either lays down arms and surrenders, or is rendered unconscious or otherwise incapacitated. The mere fact of being wounded, even to an extreme degree, does not automatically render one hors de combat.43

Second, IHL holds that "commanders and personnel should be evaluated based on *information reasonably available at the time of decision*."[44] This stipulation is known as the "Rendulic Rule," and it was established out of the judgments reached in *United States of America vs. Wilhelm List, et al.,* commonly known as "The Hostages Case." The core of this rule is that an individual "planning, authorizing, or executing military action shall only be judged on the basis of that person's assessment of the information reasonably available to the person

37 International Committee of the Red Cross, *Commentary on the Additional Protocols to the Geneva Conventions*, *supra* note 15, at 486–87, paragraph 1618.

38 International Committee of the Red Cross, *Commentary on the Additional Protocols to the Geneva Conventions*, at 487, paragraph 1619. Note that in the age of "over-the-horizon" fighting between aircraft or other vessels, visual signals will in many cases be an outdated gesture. However, crews can still communicate on open channels or utilize other clear means of signaling surrender.

39 In fact, such duplicitous conduct would constitute a perfidious act, in breach of AP I, Art. 37. Many thanks to Maciek Zając for pressing me to clarify these points more.

40 For example, certain South Korean point-defense systems are claimed to have long been capable of giving verbal warnings and demands of surrender and of recognizing multiple signs of surrender. See Simon Parkin, "Killer Robots: The Soldiers That Never Sleep," *BBC Future*, July 16, 2015, https://www.bbc.com/future/article/20150715-killer-robots-the-soldiers-that-never-sleep; Alexander Velez-Green, *The South Korean Sentry—A "Killer Robot" to Prevent War* (The Lawfare Institute, 2015).

41 Persons who are hors de combat, or "out of combat," strictly speaking cannot be considered to be combatants. For clarity and ease of exposition in what follows, however, I will refer to "combatants hors de combat" in order to distinguish these from civilians and other protected persons. See International Committee of the Red Cross, *Commentary on the Additional Protocols to the Geneva Conventions*, *supra* note 15, at 487–88, paragraph 1620.

42 AP I, Art. 41.2.c.

43 AP I, Art. 41.2.c, at 479–91.

44 Lee et al., *Law of Armed Conflict Deskbook*, *supra* note 18, at 12 (emphasis in original).

at the time the person planned, authorized, or executed the action."[45] What this means, in practice, is that different combatants will be held to different standards due to their access to differing sets of information, and, more importantly, due to how readily information even can be gathered in practice.

To return to the previous example, the infantryman clearing a trenchwork will have more direct access to clearer evidence on the status of enemy combatants. He or she is also in a position to gain clear evidence by, for example, hearing screams or looking directly at enemy combatants before firing on them. This scenario places such a combatant in a stronger position to more easily determine whether enemies should be judged to be hors de combat. Aerial combatants (or other remote fighters), on the other hand, will not have these abilities; fast-flying fighter aircraft repeatedly strafing enemy positions will have virtually no ability to determine whether initial attacks have left wounded soldiers who ought no longer to be targeted.

Under Article 15 of Geneva Convention I, combatants are required to, "without delay, take all possible measures to search for and collect the wounded and sick," a point that may seem to argue against repeatedly strafing enemy positions from the air. Yet even this seemingly strict demand "will require different kinds and levels of care" that take into account the realities of combat and the possibility for combatants to reasonably carry out such actions. The "relevant point of reference is what would be expected of a reasonable commander under the given circumstances," and "this does not mean . . . that Parties must search actively for the wounded and sick at all times, as that would be unrealistic."[46] In short, combatants are prohibited from attacking those who should be recognized to be hors de combat, and they are under a positive duty to determine that status to the extent possible, but these rules are limited by the evidence they can reasonably be expected to have, given the circumstances of their fighting.

With regard to autonomous weapons, the exact type of AWS in question will greatly impact what standards of care regulate its use, and what can reasonably be expected regarding determinations of hors de combat status. In general, an AWS's distance from the fighting and its sensor apparatuses will dictate to what extent it can recognize when someone should be deemed to be hors de combat. These factors will inform where, when, and how a combatant may permissibly deploy such systems. As optics, image recognition, and data analysis technologies improve, the ability of AWS to make reliable determinations of an enemy's wounds and attendant capacity to defend themselves will likewise improve. Even for those AWS with highly limited systems, however, distinction will not necessarily rule out their use entirely. Instead, the demands of military necessity will have to be balanced against the limits set by proportionality, which will, at most, rule out an AWS's use on a case-by-case basis.[47]

For some missions, significant risks cannot be justified, and so limited systems cannot be utilized.

The question will not be whether a system is too limited to be used at all, but rather what limitations a system has (because all systems have limitations) and what risks those limitations entail. For some missions, significant risks cannot be justified, and so limited systems cannot be utilized. But for critical endeavors where tactical failure could lead to operational or even strategic calamity,[48] it remains possible that highly limited systems may be permissibly used, despite the risks they bring.[49]

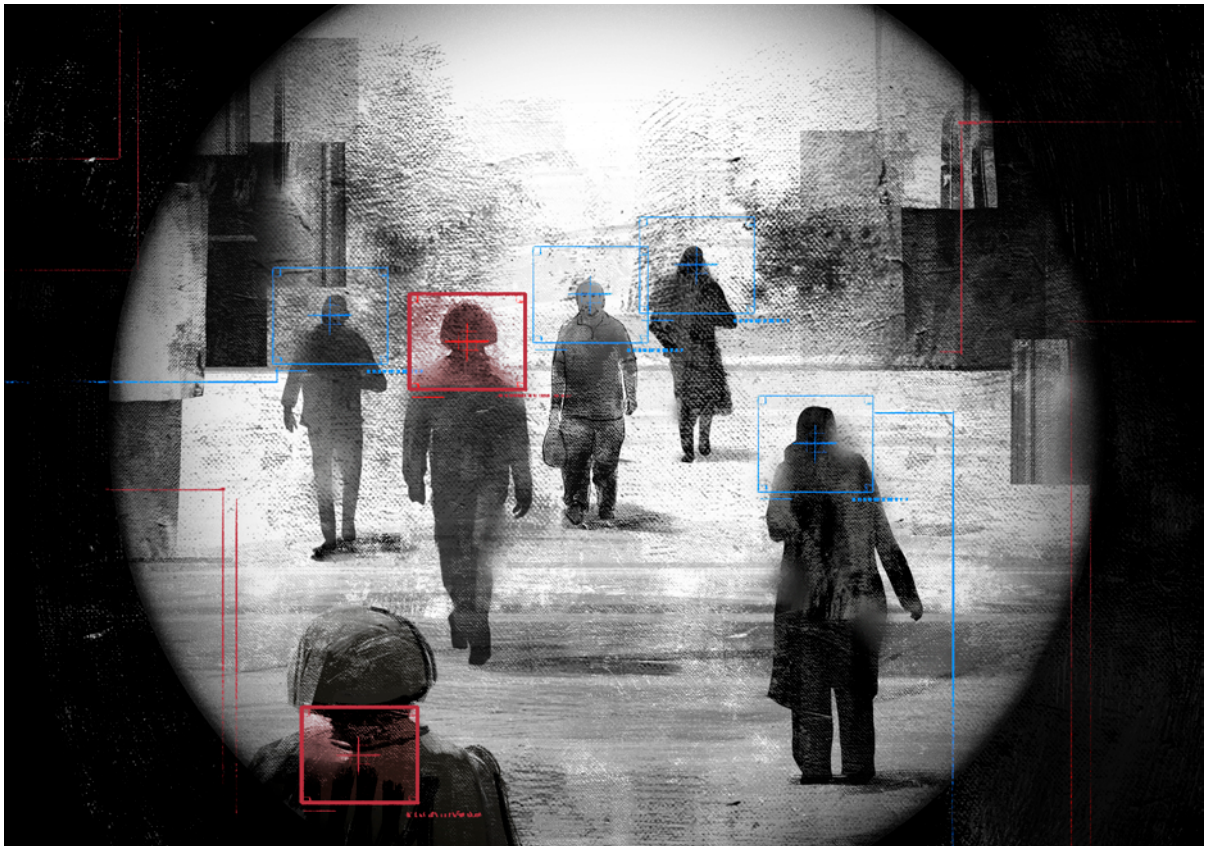
45 Lee et al., *Law of Armed Conflict Deskbook*, *supra* note 18, at 12. For further discussion of the Rendulic Rule, *The Hostages Case*, and broader elements of how evidence may impact on judgments of commanders' decisions, see Brian J. Bill, "The Rendulic 'Rule': Military Necessity, Commander's Knowledge, and Methods of Warfare," *Yearbook of International Humanitarian Law* 12 (2009), https://doi.org/10.1017/s1389135909000051; Mateusz Piątkowski, "The Rendulic Rule and the Law of the Aerial Warfare," *Polish Review of International & European Law* 2, no. 3 (2013): 69–85, https://dx.doi.org/10.2139/ssrn.2923720; Sigrid Redse Johansen, *The Military Commander's Necessity* (Cambridge University Press, 2019).

46 International Committee of the Red Cross, *Commentary on the First Geneva Convention: Convention (I) for the Amelioration of the Condition of the Wounded and Sick in Armed Forces in the Field* (Cambridge University Press, 2016), 533–34, at, respectively, paragraphs 1482, 1485, and 1486.

47 As per AP I, Art. 40, AWS that would guarantee the deaths of all enemy combatants would be impermissible, as this would amount to conducting hostilities on the basis that there will be no survivors. Thus, all AWS will arguably need to possess at least minimal means for recognizing when one is hors de combat, or must be deployed in such a manner as to accommodate this demand by default (for example, by being deployed for singular strikes where it is impossible for the AWS to kill those who have been wounded by the first attack).

48 As a historical example of this, one may consider the Allied advance into Germany and the critical importance of ensuring that at least one significant bridge across the Rhine be held intact to allow for a rapid advance of soldiers and materiel into Germany. The battle for Ludendorff Bridge at Remagen represented a case where a tactical failure would have operational and likely strategic impacts whose significance could legitimately permit the use of more destructive or less precise weapons if necessary. As proportionality limits incidental harm according to concrete and direct military advantage, massive advantage can permit increased collateral harm.

49 Critically, limitations of a system or technology do not provide a justification or even excuse that may get one "off the hook" for ethical and legal compliance. Rather, the limitations inform the ethical and legal compliance, showing combatants where and when certain systems may (not) be used. But that analysis is case specific. Thanks to the editors of this issue for pushing me to clarify this point further.



Distinction in Non-International Armed Conflict

As discussed above, non-international armed conflicts (NIAC) can be loosely characterized as warfare involving one or more parties whose fighters and military assets are not easily distinguishable from civilian persons and objects. This difference places responsibilities on combatants who might be targeted as well as those doing the targeting.

Importantly, "in order to promote the protection of the civilian population from the effects of hostilities, combatants are obliged to distinguish themselves from the civilian population while they are engaged in an attack or in a military operation preparatory to an attack" (AP I, Art. 44.3). In other words, individuals legitimately participating in combat are required to make themselves distinct in some fashion, even when they lack clear combat dress.

This requirement is a critical corollary to the core demand of distinction presented in Article 48 of AP I—namely, the act of *distinguishing* between military and civilian targets. One cannot make such distinctions unless visible markings designate relevant persons and objects as military ones.[50] International law nonetheless recognizes that in some instances, irregular combatants may be incapable of properly distinguishing themselves; in such cases, these fighters are required to carry arms openly during military engagements and when visible to enemies prior to such engagements (AP I, Art. 44.3).[51]

Whether irregular combatants are acting lawfully or not, NIAC present unique challenges of distinction for both human fighters and autonomous systems deployed to such environments. The core difficulty is the elementary act of distinguishing military persons and objects from civilian ones. This problem is exacerbated by the fact that many instances of NIAC take place in circumstances where civilians may, of necessity, also need to be regularly armed. In failed states or civil conflicts, combatants and civilians may look the same and carry the same arms. Distinction thus seems different from the object recognition exercise in IAC, and appears to present a unique difficulty for autonomous weapons systems.

50 International Committee of the Red Cross, *Commentary on the Additional Protocols to the Geneva Conventions*, *supra* note 15, at 527–29, esp. paragraph 1695.

51 Note that the requirement to carry arms openly does not necessarily imply that the arms must be easily visible. This factor may complicate easy assessment of combatant status for irregular combatants, but it further substantiates the arguments relating to AWS developed below.

This assumption, however, is too quick. It is true that civilians and irregular fighters in NIAC may be practically indistinguishable, and under international law, "in case of doubt whether a person is a civilian, that person shall be considered to be a civilian" (AP I, Art. 50.1).[52] This point would appear to undermine any lawful targeting of individuals within such conflicts. Even if we assume that every individual is actually a civilian, however, they are only protected "unless and for such time as they take a direct part in hostilities" (AP II, Art. 13.3). Thus, civilians' immunity from being objects of attack "is subject to an overriding condition, namely, on their abstaining from all hostile acts."[53] In practice, this requirement implies that civilians may be targeted in NIAC when they are directly participating in hostilities, especially through acts of violence adversely affecting the military operations of a party to the conflict.[54] This situation holds whether individuals are regularly engaged as combatants or are regularly civilians who are only currently taking a direct part in hostilities; participation in hostilities makes one a legitimate target for as long as one is so engaged.[55]

With regard to AWS, distinguishing between fighters and civilians will likely be practically impossible when no fighting is occurring. However, this inability remains the case for human combatants in such environments as well. When confronted with non-uniformed individuals lacking any distinctive signs or emblems, what distinction often demands is that one withhold fire until such individuals carry out violent actions that either threaten human combatants, friendly personnel, or civilians, or that amount to direct participation in hostilities. In other words, doubt about individuals' status is expected to be common, and such individuals should be treated as civilians unless and until such time as they directly participate in hostilities, at which point direct participation becomes grounds for lawful targeting.

One significant challenge, however, is that civilian participation is not always evident to observers, or amenable to object recognition. Serving as a lookout, calling in mortar strikes using a mobile phone, carrying supplies to combat units, and laying improvised explosive devices can all rise to the level of direct participation in hostilities, yet can be hard to detect. When in doubt, individuals are to be treated as civilians. A conservative targeting program would treat all such individuals as civilians until presented with conclusive evidence of direct participation.[56]

Despite these challenging cases, other modalities of direct participation *can* be determined largely by object recognition or algorithmic processing.[57] The key questions will often be whether an individual is armed, and if so, whether they are using their arms in a manner that permits their being targeted (for example, directing their weapons toward friendly combatants or those protected from attack).[58] Both of these questions are amenable to object recognition tasks. Moreover, targeting based on positive identification of these elements is unlikely to result in targeting individuals who would otherwise be protected from attack under IHL.[59]

It may still be the case that certain AWS cannot be deployed to specific environments, simply because they cannot carry out *these sorts* of object recognition tasks. But the conclusion then should not be that AWS in general are in breach of the principle of distinction. Rather, it should be that deploying *this* or *that* specific AWS to *this* or *that* specific operational environment would be in breach. Context will determine the legality of a system's deployment,

52 As AP I explicitly deals with international armed conflicts, one may argue that its rules do not apply in NIAC. However, the broad nature of the protections afforded to civilians, as expanded upon in the Commentary, would seem to undermine such a limitation. In particular, it is clearly stated that, "in protecting civilians against the dangers of war, the important aspect is not so much their nationality as the inoffensive character of the persons to be spared and the situation in which they find themselves" (International Committee of the Red Cross, *Commentary on the Additional Protocols to the Geneva Conventions*, *supra* note 15, at paragraph 1909). As no part of Additional Protocol II modifies or goes against this understanding, it may be taken to be authoritative with regard to NIAC, especially in light of AP II, Art. 13, which grants "general protections" to civilians in keeping with those laid out in AP I.

53 International Committee of the Red Cross, *Commentary on the Additional Protocols to the Geneva Conventions*, at 618, paragraph 1942. See also section 3 of Magdalena Pacholska, "(Il)legality of Killing Peacekeepers: The Crime of Attacking Peacekeepers in the Jurisprudence of International Criminal Tribunals," *Journal of International Criminal Justice* 13, no. 1 (2015): 43–72, https://doi.org/10.1093/jicj/mqu082.

54 For extensive discussion of the notion of direct participation in hostilities, its definition, identification, and limits, see Nils Melzer, *Interpretive Guidance on the Notion of Direct Participation in Hostilities Under International Humanitarian Law* (International Committee of the Red Cross, 2009).

55 For members of organized armed groups party to an NIAC, especially those fulfilling continuous combat functions, they may be targeted anywhere and at any time, and hostile action is unnecessary for grounding an attack. Direct participation, however, serves as a baseline for targeting when one is unsure as to the status or membership of individuals within such groups.

56 Thanks to an anonymous reviewer for reminding me of these problematic cases (for both humans and autonomous systems).

57 See tables 1 and 2 of Pacholska, "(Il)legality of Killing Peacekeepers *supra* note 53.

58 Direct participation in hostilities is much broader than what is captured by these two points, but these points represent a conservative targeting approach for AWS engaged in NIAC. Thus these "key questions" are not key questions for assessing direct participation in hostilities per se, but are only key questions for making such assessments through purely machine-operated object recognition.

59 For extensive examination of the practicalities of gaining positive identification for targeting purposes across a range of contexts, see John J. Merriam, "Affirmative Target Identification: Operationalizing the Principle of Distinction for US Warfighters," *Virginia Journal of International Law* 56, no. 1 (2016): 83–146.

and broad conclusions drawn at an abstract level are unlikely to be sustainable.

Precautions and Proportionality

AWS may present specific challenges in particular contexts, but autonomous weapons are not subject to some general objection based on the principle of distinction. Nearly every operation will have areas where autonomous weapons may be reasonably and legally deployed, as well as areas where their deployment would be almost certain to constitute a breach of IHL. If we turn our attention to the role that precautions and proportionality play in the principle of distinction, we will see that AWS deployment is not only sometimes permissible, but in certain instances required.[60]

The core of distinction is that civilians "shall not be the object of attack" (AP I, Art. 51.2). This prohibition does not imply that attacks expected to harm civilians are automatically prohibited. Rather, combatants are required to take all feasible precautions to minimize harm to civilians (AP I, Art. 57.2.a.ii), and are prohibited from carrying out attacks expected to cause civilian harms disproportionate to the concrete and direct military advantages gained (AP I, Art. 57.2.a.iii). Such limitations correspond to the *in bello* moral principles of necessity and proportionality. The latter asks, "Is this level of civilian harm acceptable, given these military outcomes?"; the former asks, "Could we reduce this harm?"

To see what distinction requires with regard to AWS and the choice of weapons and means, let us consider a hypothetical example.

> *Urban HQ: An enemy command center is located in an urban environment. Most of the personnel inside are military individuals, but there are usually between five and ten civilians as well (custodial workers, secretaries, and so forth). The command center is also adjacent to an office building that, during office hours, contains a few dozen workers. The command center is a core node of enemy command and control containing numerous high-value targets, and its destruction presents an advantage that is significant, concrete, and direct.*

The importance of the command center justifies some amount of collateral harm to civilians resulting from attacks on the building. For concreteness, let us suppose that up to twenty civilian deaths would be permissible on grounds of proportionality.[61] Given that usually no more than ten civilians are in the building, any attack destroying the whole structure would be permissible (again, purely on grounds of proportionality). Let us further suppose that the attacker has two weapons that they could deploy: a precision-guided bomb and a medium-sophistication autonomous anti-personnel drone. The AWS drone is a small quadcopter armed with a firearm and outfitted with object-recognition software capable of identifying people wearing combat dress with high reliability, and of identifying when humans are carrying weapons with medium reliability.

If use of the autonomous drone is indeed in breach of IHL (because we expect it to make these targeting mistakes), who is responsible?

If the bomb is dropped, it is expected to destroy the entire building, killing everyone inside, and posing some small risk to workers in the adjacent building (small enough that proportionality is still satisfied). The AWS, on the other hand, is expected to kill every soldier inside, by virtue of recognizing their combat dress, and is expected to target between three and five civilians as well. This outcome is expected because the AWS is known to make object recognition mistakes, such as individuals holding objects with long hafts such as mops or brooms being mistakenly identified as holding rifles, and individuals holding telephone receivers being mistakenly identified as holding pistols. These limitations make it likely that the AWS will mistakenly engage some civilians, especially janitorial and secretarial staff, because it explicitly targets them—a fact the potential deployer of the AWS knows.

Does the mistaken targeting indicate that the AWS fails distinction? One may think to argue that it does; due to a limitation in its software or hardware, the weapon identifies a civilian as a combatant, targeting that individual with lethal force. However, this conclusion mistakes where the locus of distinction

60 For useful discussion of the general challenges of satisfying proportionality when using AWS, as well as potential solutions, see, for example, Jeroen van den Boogaard, "Proportionality and Autonomous Weapons Systems," *Journal of International Humanitarian Legal Studies* 6, no. 2 (2015): 247–83, https://ssrn.com/abstract=2748997; Joseph O. Chapa, "Artificial Intelligence and the Just War Proportionality Principle," *Existenz* 17, no. 1 (2022), https://www.existenz.us/volumes/Vol.17-1Chapa.html; Maciek Zając, "AWS Compliance with the Ethical Principle of Proportionality: Three Possible Solutions," *Ethics and Information Technology* 25, no. 1 (2023): 1–13, https://doi.org/10.1007/s10676-023-09689-8.

61 The numbers assumed are purely illustrative. Attaching precise values to harms and gains is a notoriously contentious enterprise. Even so, that is part of what military lawyers must do, and some rough evaluation will usually be made in keeping with overall legal and moral norms.

lies: Distinction is first and foremost concerned with *attacks*, not weapons. In forbidding indiscriminate attacks, Article 51 of AP I does maintain that the use of methods or means of combat that cannot be directed at military objectives (or whose effects cannot be so directed) constitute indiscriminate attacks. In the above example, however, the autonomous drone is directed at a military objective: the command center. The combatant contemplating deploying the drone knows that due to limitations in its capabilities, some number of civilians are likely to be mistakenly killed. Due to limitations in the capabilities of large explosives, however, a precision-guided bomb applied to the same target is likely to kill more.

One may object that even though civilians will die when either the bomb or the AWS is used, the potential violation is not the same; the bomb does not violate distinction,[62] but may violate proportionality. By contrast, the deployment of the drone, while improving proportionality, might seem to violate distinction. The underlying objection is that when the AWS is used, civilians are not incidentally dying because they happen to be near a military objective, but are being specifically (mistakenly) targeted, and that this is a basic breach of distinction.

This objection, however, cannot be sustained. *If* use of the autonomous drone is indeed in breach of IHL (because we expect it to make these targeting mistakes), who is responsible? Who has violated IHL, when did they violate IHL, and what grouping of action and intent constituted the breach?[63] One may argue that the breach occurs at the moment of deployment—that in utilizing a weapon that cannot adequately distinguish between civilians and combatants, the combatant deploying the AWS fails in his or her duties. However, this interpretation would be incorrect; the combatant utilizes a means of attack (the drone), which can be and is directed at a military objective.[64] And if the failure of distinction is because the weapon cannot adequately distinguish between civilians and combatants, that quality is even more true of the precision-guided bomb. Bombs have absolutely no capacity to distinguish, and it is combatants who must therefore ensure that civilians are not targeted. Critically, in neither case is the combatant targeting civilians; he or she is targeting a command center, utilizing a means directed at *that target*.

One might instead argue that a breach occurs at the moment a civilian is mistakenly targeted. This objection, however, runs afoul of similar responses to those just presented. In particular, if we assume a breach occurs at that moment, who is committing the breach? The combatant has utilized a means of combat that can be directed at a military objective, which is expected to bring incidental civilian harms but ones that remain within the limits set by proportionality. Moreover, the combatant's intent is the same whether deploying the bomb or the AWS; in each instance he or she is targeting a command center, in the knowledge that civilians inside will likely be harmed.

One could shift the responsibility to the AWS itself, but this is philosophically awkward at best, and arguably simply wrong. The AWS is not highly sophisticated, and is certainly not an agent. What the AWS does is predictable for the combatant deploying it, and so can be traced back to that individual's choices and intent. This characteristic returns us to the choice a combatant faces between means of neutralizing a command center. The command center is the target, the AWS can be directed at that specific military objective, and its effects can also be so directed. Moreover, the AWS can be *more* tightly directed at that objective than the bomb can, and its effects can also be more tightly directed than the bomb's.

Any breach of distinction must be traceable to the combatant and his or her choices, but the combatant is deciding here between means of attack, both of which can be directed against a military objective. To argue that the decision to use the AWS somehow violates distinction requires reasoning that applies to the decision to use the bomb as well. However uncomfortable this fact may often be, IHL explicitly allows for incidental civilian casualties, even ones that are wholly expectable, so long as these are within the bounds set by proportionality.[65]

62 It is worth noting that some scholars, and the International Criminal Court, would actually maintain that utilizing the bomb constitutes a breach of distinction, since it is taken as a given that the only way to destroy the command center is to kill *both* the civilians and combatants. That view, however, leads to untenable and unstable ramifications, as it implies that, for example, taking involuntary human shields would render any military target untargetable. Fully exploring this debate is outside the scope of this article, but it is worth stating clearly that that view provides a perverse incentive to take hostages and locate military sites within civilian locales, thus providing a benefit to bad actors who may then exploit their adversaries' respect for the law. For that reason, I do not believe that view can be sustained, or at least, not sustained for long. Thanks to an anonymous reviewer for pressing me to clarify this further.

63 See Paola Gaeta, "Who Acts When Autonomous Weapons Strike?," *Journal of International Criminal Justice* 21, no. 5 (2024), https://doi.org/10.1093/jicj/mqae001, for detailed exploration of these questions.

64 In exploring the "scope of an attack," Jonathan Kwik makes similar points. See Jonathan Kwik, "The Scope of an Autonomous Attack," in *2024 16th International Conference on Cyber Conflict: Over the Horizon (CyCon)*, ed. C. Kwan, L. Lindström, D. Giovannelli, K. Podiņš, and D. Štrucl (IEEE, 2024), https://doi.org/10.23919/CyCon62501.2024.10685635.

65 Many thanks to Kevin Heller for pressing me to more fully explore this objection, and Jonathan Kwik for useful suggestions on clarifying the underlying points being made.

The concept of an "attack" in IHL also provides a general rebuttal to the above objection. Article 49 of AP I states that "'attacks' means acts of violence against the adversary, whether in offence or in defence" (AP I, Art. 49.1). Whether or not that violence is kinetic,[66] the concept of "attacks" presumes an *act* of violence, requiring actors or agents. Yet AWS are not actors or agents. AWS therefore do not and cannot carry out "acts." Instead, they "execute processes" on combatants' behalf.[67] As such, an AWS cannot attack (in the legal or philosophical sense) during its deployment, as it is not an entity capable of making choices or acting; it executes processes. The actors present in any deployment of a fully autonomous weapon system are the human combatants making the decisions and carrying them out. After the moment of deployment, everything that follows is the execution of a process, fundamentally akin to the way a guided missile follows its programming or an artillery shell follows the trajectory set by its initial firing position. In either of those latter cases, unexpected and unforeseeable drifts may be introduced after launch, but the missile or shell is not "deciding" where to hit. Rather, the weapon is following set programs or processes.[68]

Even for more advanced AI-enabled AWS that can select targets based on open engagement criteria or machine learning, these systems also will not rise to the level of "making decisions" or being "agentic" (in a philosophical or legal sense). Such systems are only ever executing processes—sometimes deterministically, sometimes with a stochastic spread—and though the introduction of probabilistic processes complicates the predictability of such systems, this characteristic does not make them unpredictable in a fundamental or total way. Engagement parameters, system design, or ordnance choices represent just a few of the many ways in which humans can make a system's operation more predictable, even granting some unpredictability due to AI-enabled processes.[69] Remaining unpredictability will also not present a fundamental obstacle for IHL compliance. Distinction, proportionality, and precautions in attack will tell combatants how their levels of confidence in a machine's reliability will impact on the permissibility of discrete deployments, with less predictable systems seeing fewer permissible deployments. The question is thus not whether a system is predictable, simpliciter, but whether it is *predictable enough* for a particular mission.[70]

Basic concerns over the minimization of harm further favor the AWS. Combatants are required to "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" (AP I, Art. 57.2.a.ii). Although the AWS in the above example is expected to kill *some* civilians, the bomb is expected to kill *all* civilians within the building. If the bomb's use is in accordance with proportionality considerations, the use of the AWS must be as well (since it would result in, at most, the same or a lesser number of civilian deaths, and indeed, the same civilian deaths). The AWS has the potential to be far less harmful to civilians, and is in fact expected to be; in this case, it would be *illegal* to opt for the more destructive alternative of destroying the entire building with everyone inside.[71] The presence of the adjacent office building strengthens the case, as the bomb is not expected to but nonetheless could cause some damage to the structure, while a tightly geographically limited AWS would not.

AWS will not always be superior or free of risks. An AWS not subject to temporal and geographic limitations to its operation could present grave risks to large areas, and some AWS well suited to certain tasks may be utterly unacceptable in other roles. So long as AWS are reliable in their functioning, however, and combatants can make reasonable and defensible assessments concerning what is likely to follow from certain deployments, then there will

66 For discussion of "attack" in the cyber-warfare domain, see Michael N. Schmitt, "'Attack' as a Term of Art in International Law: The Cyber Operations Context," in *2012 4th International Conference on Cyber Conflict (CYCON 2012)*.

67 See Nathan G. Wood, "Autonomous and AI-Enabled Systems: Extensions or Replacements of Human Will and Control?," *Ethics and Information Technology*, forthcoming.

68 The full arguments underpinning this point rely on involved philosophical and legal analysis that unfortunately cannot be fully presented here. My coauthors and I have further work in progress establishing these points more fully, but a similar view can already be found in Gaeta, "Who Acts When Autonomous Weapons Strike?," *supra* note 63. Many thanks to an anonymous reviewer for stressing that additional clarity was needed on this point.

69 Wood, "Autonomous and AI-Enabled Systems," *supra* note 67.

70 Similar arguments regarding the reliability of autonomous weapons can be found in Nathan G. Wood, "Reliability Standards for (Autonomous) Weapons," *Ethics and Armed Forces* 24, no. 1 (2024).

71 Importantly, if there are overriding military considerations that necessitate the drone being intact and operational for a different future mission (feasibility), then opting for use of the bomb may potentially be permissible. However, it would only be so in light of these additional considerations of military necessity. Thanks to Jonathan Kwik for suggesting this point.

be situations where the use of AWS is not just permitted, but obligatory.[72] If it is permissible to use a bomb to get the job done, then it is hard to see why a bot would be out of line.

The Realities of AWS Development and Deployment

AWS come in a variety of shapes and sizes, tailored to a variety of operational goals, and are deployed in different ways to different theaters and domains (air, sea, land, space). All of these details undermine the blanket objection that autonomous weapons cannot or will not abide by the demands of distinction. Sometimes the systems will be used indiscriminately, following every other weapon in the history of war. But they will also be used in accordance with IHL, as many already have for decades.

This brings us to a further problem in the critics' objection. When critics of AWS claim that "no autonomous robots or artificial intelligence systems have the necessary properties to enable discrimination between combatants and civilians,"[73] or that "fully autonomous weapons would not have the ability to sense or interpret the difference between soldiers and civilians, especially in contemporary combat environments,"[74] they betray a failure to grapple with the realities of AWS development and deployment.

The definition of AWS used here—weapons that "select and apply force to targets without human intervention"[75]—is one that critics also endorse.[76] This definition captures a wide variety of systems under its umbrella, from anti-personnel drones to homing torpedoes, and point-defense systems to anti-radiation missiles. The vast majority of AWS currently in use and under development are expressly anti-materiel systems: weapons and ordnance designed to destroy tanks, aircraft, missiles, ships, or other militarily relevant objects. Many of these weapon systems need not make any hard calls of distinction about whether something is a military or civilian object. Military aircraft fly far faster than civilian airliners and have IFF (Identification, Friend or Foe) tags; combat naval vessels have no civilian counterparts for which they could be reasonably mistaken; and anti-missile or active protection systems target objects that can, by definition, only be hostile.

More sophisticated weapons, even autonomous ones, do not change these responsibilities.

It is true that certain anti-materiel AWS can mistakenly target civilian objects, as anti-radiation missiles may home in on civilian radar stations, and certain anti-tank weapons may mistakenly target heavy-construction or farming equipment. These limitations, however, do not mean that such systems are "inherently indiscriminate," or that their use would automatically breach distinction. Rather, such limitations show that operators must know what risks to civilians are present in a particular environment when using particular weapons. But this is not new. When dropping bombs, firing artillery, or even firing rifle rounds, combatants are already required to take all feasible precautions to minimize risks to civilians. More sophisticated weapons, even autonomous ones, do not change these responsibilities. These obligations only demand that combatants understand their systems well enough to adequately judge the risks associated with those systems' use, and employ them accordingly.

For any autonomous weapon system, the question thus is not whether it generally can or will act in a manner conforming to distinction. Rather, for each system, combatants will have to ask themselves multiple questions: What can this AWS reliably do? What is it expected to do poorly? What are its limitations, both technical and user-imposed? Do its reliable characteristics make it likely to target any civilian or otherwise unlawful targets? If so, in which situations and why?

72 Some authors, for example, Blanchard and Taddeo, "Predictability, Distinction & Due Care in the Use of Lethal Autonomous Weapon Systems" *supra* note 4, argue that AWS will be inherently unpredictable, and as such could not be considered discriminate weapons. However, Blanchard and Taddeo's definition of AWS is much narrower than that utilized here, as it requires that AWS inherently possess forms of learning or other capabilities that would make them (eventually) unpredictable for those deploying them. Therefore this objection falls outside of the arguments presented here, as we are examining a much broader class of systems, many of which are extremely predictable. For thorough examination of their definition, see also Taddeo and Blanchard, *supra* note 7. For examination of how more futuristic AWS in line with Blanchard and Taddeo's conception may be handled under IHL, see Alan L. Schuller, "At the Crossroads of Control: The Intersection of Artificial Intelligence in Autonomous Weapon Systems with International Humanitarian Law," *Harvard National Security Journal* 8, no. 2 (2017): 379–425, https://ssrn.com/abstract=2978141; Tim McFarland, "Autonomous Weapon Systems and IHL Compliance: A Constrained Legal Optimisation Problem," *Journal of Military Studies* 1, no. 1 (2023), https://doi.org/10.2478/jms-2023-0006.

73 Sharkey, "Saying 'No!' to Lethal Autonomous Targeting," *supra* note 3, at 378.

74 Human Rights Watch, *Losing Humanity*, *supra* note 3, at 30.

75 International Committee of the Red Cross, *ICRC Position on Autonomous Weapons Systems*, *supra* note 8, at 1.

76 See note 12.

The answers to these questions will not tell the combatant whether the AWS "satisfies distinction," as weapons do not do this on their own. Instead, these answers will inform him or her about what particular uses or deployments of the AWS would be permitted on grounds of distinction. Put negatively, distinction, along with a clear eye to the limitations and expected failings of a particular autonomous system, will tell which uses or deployments would be indiscriminate and thus illegal.

Distinction does not decree that all high-yield bombs are "indiscriminate" by virtue of having massively destructive potential; it only tells us that certain uses of those weapons are indiscriminate. In the same way, AWS, even very primitive ones with little ability to distinguish combatants from civilians, will not be outlawed simply due to their limitations. Rather, the limitations in capability place limitations on the set of deployments legally permissible for those weapons. It is still the responsibility of men and women in military organizations to ensure that AWS are deployed according to their capabilities. Failure to do so is not a failure of the autonomous system; it is a failure of choosing the wrong weapon for a specific task in war.

Conclusion

A rifle does not discriminate. Nor does a hand grenade, an artillery shell, or a bomb. And importantly, they are not required to—IHL imposes obligations on belligerents, not their equipment. Distinguishing between legitimate and illegitimate targets is something that the combatants behind weapons do, and have done.

Yet we are now moving into an age where weapons increasingly can make distinctions. Sometimes these weapons will fail, and sometimes there will be situations where they cannot be expected to succeed. Even so, there are a host of cases where specific autonomous systems, despite their many limitations, can reliably distinguish between military targets and protected persons and objects, and in some cases already are doing so. The fact that AWS sometimes fail in this is not a mark against them as a category of weapon systems. Rather, the fact that they ever succeed is a significant achievement for both humanity and our hopes of a less bloody future of war. We therefore have a responsibility to ensure that deployments of AWS are conducted in accordance with the requirements of distinction—but this is no different from what is required when we deploy any other weapon, be it firing a missile, dropping a bomb, tossing a grenade, or aiming a rifle.

While certain deployments of AWS may satisfy distinction, others clearly will not, and the principle of distinction will thus outlaw these. Again, however, this is nothing novel; all weapons of war may be used in more or less discriminate ways. The capabilities and limitations of each individual autonomous system will determine what precautions and standards of care are required, and what sorts of deployment are permissible. In some cases, certain AWS will be inappropriate, but in other situations, autonomous weapons are likely to be superior, both morally and legally, to conventional alternatives. The context of an engagement is what determines which weapons may or may not be used. Weapons themselves are only legal, or not, in virtue of the evidence and information available to combatants who must decide how to proceed in war.

The International Committee of the Red Cross reminds us that, "unable to eliminate the scourge of war, one endeavors to master it and mitigate its effects."[77] Autonomous weapon systems bring many risks and concerns, but they also present an opportunity to, in specific cases, wage more discriminate forms of war—an opportunity to master and mitigate the scourges of war. We must act with caution and care as we explore the potential of these types of weapons, but we ought also to remember the possibilities they bring.

Nathan G. Wood *is the junior research group leader for the "MilEth" project (Military Defense Technologies and Ethics) at the Institute of Air Transportation Systems of the Hamburg University of Technology. He is also an associate member of the Center for Environmental and Technology Ethics–Prague, and an external fellow of the Ethics + Emerging Sciences Group at California Polytechnic State University San Luis Obispo. His research focuses on the ethics and laws of war, especially as these relate to emerging technologies, autonomous weapon systems, outer space warfare, and other aspects of future conflict. He has previously published in* Ethics and Information Technology, War on the Rocks, Philosophical Studies, The Journal of Military Ethics*, and numerous other journals.*

Acknowledgments: *This article delves into deep questions of law and philosophy, and I am grateful to Isaac Taylor and Magdalena Pacholska for insightful feedback on both those fronts. I must also thank Kevin Jon Heller and Lena Trabucco for suggestions on a core (and controversial) contention I make relating to IHL; their feedback helped*

77 International Committee of the Red Cross, *Commentary on the Additional Protocols to the Geneva Conventions*, *supra* note 15, at 480, paragraph 1610.

iron out the presentation to a great degree. My sincerest thanks go also to Jonathan Kwik and Maciek Zając, both of whom gave detailed and constructive criticism to the piece at critical stages. Finally, I must thank the two anonymous reviewers of this article and the entire editorial team at Texas National Security Review *for numerous helpful suggestions and assistance in cleaning up the final manuscript for publication.*

Funding: *This work was partially supported by the Grantová agentura České republiky (Czech Science Foundation), under grant #24-12638I, and by the "MilEth" (Military Defense Technologies and Ethics) research project, funded by the Bundesministerium für Forschung, Technologie und Raumfahrt (German Federal Ministry of Research, Technology, and Space) under grant #01UG2409A.*

Image: *US Army Transformation and Training Command by Luke Allen.*[78]

78 For image, see https://www.dvidshub.net/image/6153037/next-generation-combat-vehicle-cross-functional-team-demonstrations.