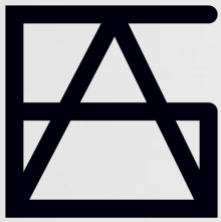
June 2024



ELSA LAB
DEFENCE

# No. 4

# Research paper series

## Scope of an Autonomous Attack

AUTHOR
Jonathan Kwik

**CyCon 2024: Over the Horizon**
**16th International Conference on Cyber Conflict**
C. Kwan, L. Lindström, D. Giovannelli, K. Podiņš, D. Štrucl


# The Scope of an Autonomous Attack

**Jonathan Kwik**
Postdoctoral researcher
T.M.C. Asser Institute
The Hague, Netherlands
j.kwik@asser.nl; j.h.c.kwik@gmail.com

**Abstract:** 'Attack' is an important term of art in international humanitarian law that serves as the basic unit of reference for many targeting obligations. It is often also asserted that human commanders of autonomous weapon systems (AWS) must make legal determinations 'per individual attack'. Divergent interpretations on what constitutes an attack nevertheless lead to drastically different conclusions with regard to the technology's lawfulness: interpreted narrowly ('each shot'), it precludesAWS technology entirely, while interpreted broadly ('each activation'), it sanctions extensive autonomous activity. This paper theorises that *imprecision on the scope of attack* is an underappreciated aspect of the AWS controversy that hampers theoretical and diplomatic advancements. The legal boundaries of autonomous attacks are analysed through the lens of targeting law, and a scaling methodology is proposed that allows commanders to determine the maximum extent to which autonomous activity may still lawfully be grouped into one single attack. The paper argues that both overly narrow and broad interpretations are inconsistent with targeting principles and practice, instead favouring a middle-ground approach based on temporal and spatial proximity that properly respects international humanitarian law's (IHL) balancing philosophy between humanitarian and military interests. Through consideration of practical scenarios, the paper subsequently demonstrates how this impacts the application of targeting rules, such as at what intervals the commander's duty to verify or cancel is triggered and under what circumstances successive autonomous engagements may be grouped together for proportionality assessments.

**Keywords:** *attack, targeting, autonomous weapons, precautions, IHL proportionality, proximity in time and space*

# 1. INTRODUCTION

'Attack', defined in Additional Protocol I (API) Article 49(1) as 'acts of violence against the adversary, whether in offence or in defence', is an important term of art in international humanitarian law (IHL) to which many protections attach.[1] New technologies sometimes necessitate revisiting how 'attack' is interpreted. In the case of cyberweapons, their intangible nature provoked a shift from the traditional (physical) conception of *violence* to a more effects-based approach.[2] In contrast, the term has received less academic attention in the debate surrounding autonomous weapon systems (AWS). While there has been extensive discussion on whether AWS can be used to lawfully conduct attacks or can be designed to properly implement precautions,[3] it is usually presumed that the 'attack' component is relatively uncontroversial. The infliction of violence is an attack's *sine qua non*,[4] and as AWS are usually conceived as physical systems intended to inflict (physical) harm to military objectives,[5] there is little reason to doubt that employing an AWS constitutes an attack. It is also uncontended that commanders employing AWS are obligated to ensure that fundamental principles such as distinction, precautions and proportionality are upheld.[6]

However, there is a different and underappreciated point of legal uncertainty regarding the notion of attack as it relates to AWS – one that significantly impacts how targeting rules are applied to AWS attacks.

Consider the following scenario:

> **Scenario 1.** At 1200, Commander-A activated an AWS to attack a tank platoon, during which the system released seven shots at four tanks (successive shots were released because the AWS detected that the objective

1 MN Schmitt, '"Attack" as a Term of Art in International Law: The Cyber Operations Context' in C Czosseck, R Ottis and K Ziolkowski (eds), *4th International Conference on Cyber Conflict* (NATO CCDCOE 2012) 284.

2 See C Droege, 'Get off My Cloud: Cyber Warfare, International Humanitarian Law, and the Protection of Civilians' (2012) 94 International Review of the Red Cross 533, 552–557.

3 See eg WH Boothby, 'Highly Automated and Autonomous Technologies' in WH Boothby (ed), *New Technologies and the Law in War and Peace* (Cambridge University Press 2018).

4 Schmitt (n 1) 290.

5 Eg N Davison, 'A Legal Perspective: Autonomous Weapon Systems under International Humanitarian Law', *UNODA Occasional Papers No. 30* (2017) 5. There is no consensus on how to define AWS. The term is used in this paper to refer to physical weapon systems enabled by artificial intelligence that can execute target selection and engagement independently. Whether this process actually occurs without human oversight is an operational choice made by the deploying commander.

6 There is universal consensus that targeting rules remain applicable to the use of AWS. See GGE on LAWS, 'Report of the 2023 Session of the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems' (24 May 2023) CCW/GGE.1/2023/2 para 21(a); M Pacholska, 'Autonomous Weapons' in Bartosz Brożek, Olia Kanevskaia and Przemysław Pałka (eds), *Research Handbook on Law and Technology* (Edward Elgar Publishing 2023).

was still operational after the initial strike). At 1500, the system was sent to attack four tanks guarding the city's four arteries, during which the AWS released six shots.

How many attacks has Commander-A launched today? That depends on the unit of measurement for 'attack':

- Each activation 2 attacks
- Each tank 8 attacks
- Each shot 13 attacks

This question is fundamental as 'attack' serves as the basic unit of reference for many targeting obligations.[7] 'Attack' is also used as a yardstick in many policy and diplomatic proposals. Take the proposition that a 'human should be in control of the system for each individual attack'.[8] This statement is actually not particularly controversial: precautionary obligations are addressed at 'those who plan or decide upon an attack',[9] and there is relatively broad international agreement that a machine cannot discharge legal obligations 'for' its human user (i.e., the commander).[10] However, depending on how broadly or narrowly one defines 'attack', this same proposition implies entirely divergent requirements for human involvement. At one extreme, we find interpretations that construe 'attack' at the narrowest level, such as at 'the stage when the munition is fired'.[11] This essentially 'precludes autonomous systems' altogether.[12] At the other end of the spectrum, we find positions arguing that an attack could potentially encompass the entire period in which the system is active (during which it may release multiple shots at many different objectives).[13] In concept, this permits autonomous operation at the 'trigger-pulling' level.[14]

The author of this paper perceives much of the disagreement in literature and debate regarding the lawfulness of autonomous technologies as derived from *imprecision and diverging interpretations on the scope of attack*. The aims of this paper are to explore how the scope of an AWS attack should be conceptualised legally based

7 HM Roff and R Moyes, 'Meaningful Human Control, Artificial Intelligence and Autonomous Weapons' (2016) Briefing Paper for the Informal Meeting of Experts on Lethal Autonomous Weapons Systems, 5.

8 T Chengeta, 'Defining the Emerging Notion of "Meaningful Human Control" in Autonomous Weapon Systems' (2016) 49 International Law and Politics 833, 875.

9 Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts (adopted 8 June 1977, entered into force 7 December 1978) 1125 UNTS 3 ['API'] art 57(2)(a).

10 GGE (n 6) para 21(c); ICRC, 'Ethics and Autonomous Weapon Systems: An Ethical Basis for Human Control?' (2018) CCW/GGE1/2018/WP, para 32; US Department of Defense, *Law of War Manual, Updated July 2023* (US Department of Defense 2015) art 6.5.9.3.

11 W Boothby, 'Control in Weapons Law' in Rogier Bartels and others (eds), *Military Operations and the Notion of Control Under International Law* (TMC Asser Press 2021) 388.

12 ET Jensen, 'Autonomy and Precautions in the Law of Armed Conflict' in Rain Liivoja and Ann Väljataga (eds), *Autonomous Cyber Capabilities under International Law* (NATO CCDCOE 2021) 191.

13 Eg M Ekelhof, 'Moving Beyond Semantics on Autonomous Weapons: Meaningful Human Control in Operation' (2019) 10 Global Policy 343.

14 That is, each time a munition is released. See Jensen (n 12) 192.

on currently applicable targeting law and to propose a reasoned methodology to determine the appropriate scope of specific operational scenarios. Greater clarity on this legal question is beneficial both theoretically and practically. In terms of theory, it advances the doctrinal debate by offering greater precision as to the operational circumstances within which AWS would conflict with targeting requirements, such as proportionality and precautions. In terms of practice, the methodology provides field commanders with a readily usable cognitive framework with which to consistently determine how broadly or narrowly they may define 'attack' when AWS are planned for use during operations.

With this background established, the paper proceeds as follows. First, Section 2 elucidates the notion of *scope* as it relates to attacks under IHL and proposes a qualitative scale that allows us to theorise on the possible ways *Attack Scope* can be conceived. With this as the starting point, Section 3 considers the contours of how broadly or narrowly IHL permits attacks to be construed. It is argued that both overly narrow and overly broad conceptions of attack run counter to the purposes of IHL. Instead, a middle ground based on temporal and spatial criteria is recommended, which properly balances the humanitarian goals of IHL with the practicality of targeting. Section 4 applies this theory to targeting scenarios to demonstrate how commanders can use the proposed methodology to assess the appropriate scope of AWS attacks and how this impacts obligations in *attack* (in particular targeting rules in API Arts. 51 and 57). It also discusses the additional legal insights that this theorisation provides. Section 5 concludes with overall remarks and recommendations.

Note that 'the commander, not the weapon system, makes legal determinations'.[15] This paper assumes that legal or moral agency cannot be assigned to machines. They can *perform functions consistent with* requirements such as distinction and proportionality,[16] but they cannot discharge IHL 'for' the human commander.[17] As such, the commander remains the primary person responsible for implementing precautionary obligations, and the steps proposed below are intended to be applied by a human commander prior to launching an AWS attack. This paper focuses on how the *user* (the commander) can ensure that obligations in *attack* remain respected when AWS are involved.

15 RJ Sleesman and TC Huntley, 'Lethal Autonomous Weapon Systems: An Overview' (2019) 1 Army Lawyer 32, 34. In making this determination, the commander may consult with a legal adviser, who helps them make a reasoned decision based on applicable law, the operational circumstances, and a risk assessment. Ultimately, however, the legal determination and responsibility remain the commander's. TD Gill and D Fleck (eds), *The Handbook of the International Law of Military Operations* (Oxford University Press 2010) para 31.02.

16 As such, no position is taken with respect to design requirements sometimes raised in literature, such as whether an AWS must be able to autonomously calculate proportionality or cancel attacks.

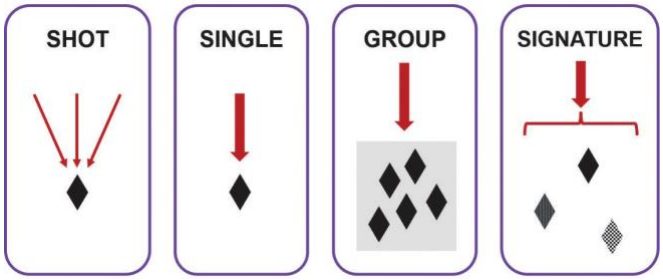
17 Above n 10.

## 2. DIFFERENT SCOPES OF ATTACKS

First, let us clarify what is meant by *Attack Scope*. An attack has been variably described as 'combat action',[18] 'the commission of acts of violence',[19] or 'any military act of a violent nature'.[20] As indicated by the last quote, the threshold is verylow: even a single sniper shot or lone dropped bomb qualifies as an attack.[21] There isalso a clear limit at the other end of the spectrum: in any event, an attack is narrowerthan an operation.[22] Between these two extremes, however, there is less clarity on how one should determine the Attack Scope of a particular use-of-force instance, as Scenario 1 showed.

Figure 1 provides a visualisation of the possible ways the scope of an AWS attack could be characterised.

**FIGURE 1:** ATTACK SCOPE SPECTRUM




*Shot* is the narrowest way Attack Scope can be construed. Here, each 'trigger-pulling action' by the AWS (i.e., each bullet shot or munition released) is considered a separate attack, even if these are successively aimed at the same objective. In contrast, *Single* considers all shots taken against the same target entity[23] as one attack.

Another possibility is to allow strikes against multiple entities to be combined. Under *Group*, all strikes against a Specific Target Group may be combined for legal purposes. The US Department of Defense defines a Specific Target Group as a 'discrete group of potential targets, such as a particular flight of enemy aircraft, a particular formation

18 Y Sandoz, C Swinarski and B Zimmerman, *Commentary on the Additional Protocols of 8 June 1977 to the Geneva Conventions of 12 August 1949* (Martinus Nijhoff 1987) ['AP Commentary'] para 1880.

19 *Prosecutor v Kunarac, Kovać and Vuković* (ICTY, Trial Judgment) IT-96-23-T & IT-96-23/1-T (22 February 2001) para 415.

20 Program on HPCR at Harvard University, *HPCR Manual on International Law Applicable to Air and Missile Warfare* (Cambridge University Press 2013) ['HPCR Manual'] 28.

21 M Bothe, KJ Partsch and WA Solf (eds), *New Rules for Victims of Armed Conflict: Commentary on the Two 1977 Protocols Additional to the Geneva Conventions of 1949* (Martinus Nijhoff 2013) ['2013 Commentary'] 329.

22 D Fleck (ed), *The Handbook of International Humanitarian Law* (Oxford University Press 2013) para 442.

23 'Entity' can refer to either a person or an object.

of enemy tanks, or a particular flotilla of enemy vessels'.[24] Therefore, to qualify as a Specific Target Group, the entities must share some proximity in both time and space. In Scenario 1, the platoon of four tanks targeted at 1200 hours can be considered a Specific Target Group. If characterised as a Group-level attack, all seven munitions released by Commander-A's AWS can thus be considered as a single attack.

The broadest approach is to consider all engagements against objectives with a shared *Signature*[25] to be part of a single attack. For example, a commander may activate an AWS to attack 'a particular model of tank or aircraft' within a specific area.[26] Compared to Group, the difference lies in the fact that these entities are not part of a *discrete* group as mentioned above. Many current in-use systems fall into this tier. 'Defensive' systems such as C-RAM, ground-based sentries and air defence use signatures such as trajectory, form, size, etc. to determine whether entities fall within the designated threat profile.[27] The Tomahawk, in so far it was launched over the horizon against a presumed presence of Soviet ships (but not any particular one),[28] would also qualify. As Scharre remarks with regard to the Harpy, its user 'does not know […] *which particular* radars are to be engaged, only that radars that meet the Harpy's programmed parameters will be'.[29]

Using this scaling system allows us to analyse any use-of-force situation and establish the boundaries of where each attack legally ends and a new attack begins. To return to Scenario 1, *Shot* would consider that the AWS is engaging in a separate attack each time it detects an operational tank and releases a munition, even in iterative situations against tanks it previously engaged and failed to disable. Given that (human) commanders are required to engage in legal analysis prior to each attack(on the military nature of the objective and the proportionality of the engagement),[30] and granting that IHL 'precludes AWS from moving from one "attack" to another

24 US Department of Defense, 'Autonomy in Weapon Systems' (2023) DoD Directive 3000.09 ['DoD Autonomy'] 23.

25 A signature is a 'pattern of sensor data that is taken to represent a target'. R Moyes, 'Target Profiles: An Initial Consideration of "Target Profiles" as a Basis for Rule-Making in the Context of Discussions on Autonomy in Weapons Systems' (Article 36 2019) 4. Landmines use weight to "decide" whether to explode. Modern AI systems can use complex signature combinations (eg heat, smoke, contextual semantic information) to determine whether an entity falls within the intended target set. See eg F Meng and others, 'Visual-Simulation Region Proposal and Generative Adversarial Network Based Ground Military Target Recognition' (2022) 18 Defence Technology 2083.

26 DoD Autonomy (n 24) 23.

27 M Ekelhof, 'Lifting the Fog of Targeting: "Autonomous Weapons" and Human Control through the Lens of Military Targeting' (2018) 71 Naval War College Review 61, 74.

28 See J Markoff, 'Fearing Bombs That Can Pick Whom to Kill' *New York Times* (11 November 2014) <www.nytimes.com/2014/11/12/science/weapons-directed-by-robots-not-humans-raise-ethical-questions.html> accessed 2 August 2023.

29 PD Scharre, 'Autonomy, "Killer Robots," and Human Control in the Use of Force' (*Just Security*, 9 July 2014) <https://www.justsecurity.org/12708/autonomy-killer-robots-human-control-force-part/> accessed 10 June 2021 (emphasis original).

30 A Cohen and D Zlotogorski, *Proportionality in International Humanitarian Law* (Oxford University Press 2021) 65.

[…] without each individual attack being subject to human legal judgments',[31] this imposes a strict obligation on commanders to continuously monitor their AWS (and effectively precludes the lawful use of autonomous systems altogether).[32]

*Single* is less demanding. Consider the attack on the tank guarding the south artery at 1500 hours. Under this interpretation, the commander is required to confirm the military nature of the tank and proportionality *prior* to commencing their AWS attack (i.e., activating the system)[33] but is not required to repeat this assessment every time the AWS detects that another strike is necessary to disable the tank: they can leave the AWS to 'finish the job' independently and still be confident that they properly discharged all duties in attack. Compared to Shot, Single introduces slightly more risk with respect to the protections offered by IHL. For instance, what if, between shots, the tank unexpectedly rolls next to a market stall in an attempt to evade the AWS?[34] Improperly delineated, Single can also be taken to an absurd extreme. Suppose the commander orders an insurgent leader to be targeted based on biometric signatures, but the leader manages to dodge an initial attempt by the loitering system. Six hours later, he resurfaces again and is identified and killed by the AWS. Is this still part of the same attack? If not, at what point was the commander required to conduct their legal evaluations again?

*Group* and *Signature* are even more militarily efficient. For the 1200 sortie, construing the tanks as one group allows the commander to perform validation and the proportionality assessment only once for the platoon as a whole before activating the AWS. With regard to proportionality, imagine that for the 1500 sortie, the commander learns that attacking the north tank will unavoidably hit a market stall. Combining all four tanks for the purposes of military advantage might allow the commander to proceed with the attack on the north tank, compared to if this were classified as four Single-attacks.[35] Conversely, these broader conceptions of Attack Scope further amplify the risk to the civilian population, both by increasing the epistemic distance between the commander and the effects of the attack, and by potentially justifying significant magnitudes of collateral damage under the pretence that all incidental harm was inflicted 'under the same attack'.[36] The question thus arises: are there limitsto how broadly or narrowly Attack Scope can be construed?

31 Roff and Moyes (n 7) 5.
32 Jensen (n 12) 191.
33 API (n 9) art 57(2)(a)(i) and (iii) respectively.
34 A fighter pilot in a similar situation would likely delay their re-strike for fear of excessive collateral damage.
35 See Section 4.
36 See Section 3.

## 3. DELINEATING ATTACK SCOPE

Moving between Attack Scope tiers involves trade-offs in protection versus practicality. The lowest tier (Shot) offers the greatest level of safety to the civilian population but is very demanding from a military perspective. It not only requires constant human supervision and legal evaluation per shot[37] but also construes the proportionality rule very narrowly (this limitation also applies to Single). Moving up the scale to Group and Signature, we see the opposite effect as efficiency becomes the dominant factor. Group allows commanders the benefit of only performing validation and proportionality evaluations once for a collective of objectives. Signature is yet more permissive since the objectives need not even constitute a disparate group. We can thus hold that from a military perspective, construing Attack Scope broadly is preferred. Those prioritising military efficiency will want to define the scope of their AWS attack as broadly as possible, as this confers practical and logistical benefits. Those emphasising humanitarian interests will want to take the opposite position, and push for AWS attacks to be construed as narrowly as possible to maximise civilian protection.

Given these competing interests and the doctrinal ambiguity on this matter, there is clear value in establishing clear and reasoned guidance as to how far commanders may stretch the notion of *attack* when planning use-of-force using an AWS. To this end, this section first explores the theoretical contours that limit how narrowly or broadly Attack Scope *can* be conceived. Then, having identified these contours, more concrete criteria are provided, which determine to what extent a commander *may* aggregate multiple strikes/objectives into one AWS attack.

### *Contours*

Drawing the line too far to the left or right of the scale is conceptually problematic in either direction. To see why, take the most restrictive option of always construing *each shot* as a separate attack. As noted above, many positions in the AWS polemics imply this proposition, but this is conceptually incorrect, inconsistent with practice and makes many related rules unworkable. Both API and the International Criminal Tribunal for the former Yugoslavia (ICTY) refer to '*acts* of violence'.[38] The API Commentary similarly mentions 'combat action' and 'counter-*attacks*',[39] implying that several individual strikes may agglomerate into one attack.[40] Bothe et al. confirm that a Shot-only interpretation was never the API drafters' intent.[41]

37 This is not only a logistical burden, but also precludes technologies offering benefits that cannot be achieved at human levels of performance, for example, when the time of engagement would be too short for a human response.

38 API (n 9) art 49(1); *Kunarac* (n 19) para 415 (emphasis added).

39 AP Commentary (n 18) para 1880 (emphasis added).

40 2013 Commentary (n 21) 329.

41 ibid.

Additionally, the proposition is inconsistent with existing practice. An attack may involve the use of force against multiple objectives that do not permit legal scrutiny per engagement.[42] Many 'defensive' systems currently in use, such as C-RAM and active protection systems,[43] would be impossible to operate if each shot were classified as an attack for which a separate legal analysis must be conducted.[44] Commanders do not supervise when each landmine 'decides' to release force or not.[45] For rocket volleys fired at a concentration of tanks, one does not look at each individual munition and determine whether it was properly directed at a military objective, whether collateral damage is excessive, etc.:[46] instead, the volley as a whole is assessed. In each of these examples, the legal analysis is performed *once*, prior to the attack's commencement.

Finally, the proportionality rule – which requires that expected collateral damage may not be excessive relative to the anticipated military advantage – precludes a Shot-only interpretation. With respect to proportionality evaluations, the possibility of counting several strikes together for the purposes of determining an attack's military advantage has been debated in the past. Commentators considered whether military advantage should be assessed on the basis of 'a single strike or for a series of strikes',[47] 'a single military objective, on the basis of a battle, a campaign or a war'.[48] Consensus was eventually reached in both scholarship and practice[49] that military advantage should 'relate to the attack considered as whole and not merely to isolated or particular parts of the attack',[50] indicating that an attack may indeed consist of multiple strikes on one or more entities.[51]

At the same time, defining 'attack' too broadly is also contrary to existing law and threatens many protections accorded by IHL. Take the proposition that '*each activation* is an attack'.[52] Depending on how long the AWS is allowed to remain active, its freedom of movement, how specific the target signature is, etc., this 'one attack' can theoretically span hours, many square kilometres, hundreds of objectives, and thousands of shots. Many protections granted by distinction and precautions would be compromised if commanders were permitted to amalgamate all this activity

42 Roff and Moyes (n 7) 5.

43 Many have been in use for several decades. See P Scharre and MC Horowitz, 'An Introduction to Autonomy in Weapon Systems' (2015) Center for a New American Security, Annex B.

44 Scharre (n 29).

45 See Moyes (n 25) 4.

46 Assuming proper corrections are made after initial shots.

47 N Durhin, 'Protecting Civilians in Urban Areas: A Military Perspective on the Application of International Humanitarian Law' (2016) 98 International Review of the Red Cross 177, 188.

48 WJ Fenrick, 'International Humanitarian Law and Combat Casualties' (2005) 21 Revue européenne de Démographie 167, 177

49 See US Department of Defense (n 10) sct 5.6.7.3; Fleck (n 22) para 445; 2013 Commentary (n 21) 366.

50 HPCR Manual (n 20) para 1(w).

51 Y Dinstein, *The Conduct of Hostilities under the Law of International Armed Conflict* (Cambridge University Press 2016) 108.

52 Cf Scenario 1, in which case the 1200 and 1500 sorties would count as one attack respectively, for a total of two attacks.

into one attack.[53] Conceptually, it would also render the proportionality rule moot. 'Attack' can in no event stretch to encompass strikes conducted during an entire military operation[54] since this would justify almost endless levels of collateral damage on the basis that it permitted the belligerent to 'win the battle' or even the entire war.[55]

We can thus generally conclude that commanders must be allowed to conceive their AWS attacks sufficiently broadly to allow some engagements to be taken by the system in succession without human legal judgment being required for each, yet narrowly enough so as not to jeopardise the protections afforded by IHL. The question that then follows is: how do we determine the appropriate scope for *specific* AWS attacks?

## *Criteria Limiting Permissible Scope*

An attack has to 'remain a finite operation with defined limits'.[56] This paper argues that two factors should inform the maximum extent to which a commander may consider successive engagements by their AWS as one attack. These are *proximity in time and proximity in space*.[57]

In API's *travaux préparatoires*, the International Committee of the Red Cross (ICRC) remarked that an attack 'is related to only one specific military operation, limited in space and time'.[58] Similarly, Dinstein argued that 'the temporal or geographic dimensions must be construed reasonably. They cannot be too remote or long-term.'[59] The ICTY commission investigating NATO conduct during the Kosovo War also invoked 'time or space' as a key delimiter for assessing the proportionality of particular attacks.[60] With regard to spatial proximity in particular, recall that API considers any attack that 'treats as a single military objective a number of *clearly separated* and distinct military objectives' as indiscriminate.[61] This presumption that time and space are decisive is reflected in works on AWS. Many commentators emphasise the need for restricting the time and space in which an AWS will operate,[62] and the Group of Governmental Experts in Geneva held that AWS users should '[l]imit the *duration, geographical scope,* and *scale* of the operation of the weapon system'.[63]

53 Considered from this perspective, the assertion that an AWS 'cannot proceed from one attack to another, to another, without human legal judgment being applied' makes eminent sense. See Article 36, 'Key Elements of Meaningful Human Control, Background Paper to Comments Prepared by Richard Moyes, Managing Partner, Article 36' (2016) CCW Meeting of Experts on LAWS, Geneva, 11–15 April 2016, 3.

54 Fleck (n 22) para 442.

55 L Gisel, 'The Principle of Proportionality in the Rules Governing the Conduct of Hostilities under International Humanitarian Law', (International Expert Meeting, 22–23 June 2016, ICRC 2016) 13.

56 ibid 17.

57 'Space' may need to be defined differently in the case of systems that attack intangible objectives, which fall outside the scope of this paper.

58 2013 Commentary (n 21) 329 (fn 2).

59 Dinstein (n 51) 161.

60 ICTY, 'Final Report to the Prosecutor by the Committee Established to Review the NATO Bombing Campaign Against the Federal Republic of Yugoslavia' (2001) para 49.

61 API (n 9) art 51(5)(a) (emphasis added).

62 Eg Boothby (n 3) 145; P Kalmanovitz, 'Judgment, Liability and the Risks of Riskless Warfare' in Nehal Bhuta and others (eds), *Autonomous Weapons Systems* (Cambridge University Press 2016) 150.

63 GGE (n 6) para 22.

In particular, an absolute line must be drawn at the point where spatial and temporal distance starts to compromise a commander's ability to properly implement precautions in attack due to epistemic uncertainties. Since the commander

> needs to make legal judgements based on an anticipation of the interaction of a system with its operational context, there needs to be some bounding of that context in space and time in order for such judgements to be substantive. The wider the physical area, and the longer the duration of operation, the less detailed the information a commander will likely have regarding that area, and the less predictable that system's use will be.[64]

How dynamic the operational environment is, as well as the density and type of collateral concerns, influence this assessment. Some environments are relatively static or predictable (e.g., the deep sea, a demilitarised zone), but other environments in which AWS usage may be envisaged (e.g., populated areas) can change quickly and unexpectedly.[65] In complex and dynamic contexts, the maximum permitted timeframe within which objectives may still reasonably be grouped into one attack will likely be very limited, particularly if many (mobile) collateral concerns are present. Spatially, objectives grouped together in close proximity can more easily be assessed on the basis of similar surrounding circumstances, while objectives spaced further apart will inhabit distinct operational spaces that cannot be legally analysed as one attack. Finally, the commander should also consider the expected delay between engagements. There is clearly a distinction between sending multiple AWS to attack objectives simultaneously and relying on only a few (or one) AWS to locate and strike targets, since targets currently *not* being engaged may be free to move, making prior collateral estimations obsolete.

These judgments are context-specific, and commanders must carefully analyse all operational and environmental parameters to determine whether they can, in good faith, order a particular AWS attack and be reasonably convinced that all presumptions relevant for legal analysis (concerning the military nature of the objective, military advantage, collateral damage, etc.) *hold throughout* the planned autonomous attack.

One may argue that the guidelines proposed above are not specific enough or even that they are open to abuse. Would it not be better to establish more quantitative standards in terms of seconds or square kilometres? This paper argues in the negative. The extent to which objectives 'should be geographically proximate to each other, and theduration over which a use of force may constitute an individual attack, are all open questions to some extent'.[66] The aforementioned ICTY report also declined to give rigid guidelines, preferring instead to leave the question open.[67] The current author

64 Moyes (n 25) 9.
65 ICRC (n 10) para 43.
66 Roff and Moyes (n 7) 5.
67 ICTY (n 60) para 49.

views this flexibility as desirable. IHL is a practical legal regime that recognises that each targeting situation is unique.[68] Concepts such as maximum permissible Attack Scope should, therefore, not be reduced to mathematical standards: as with precautionary obligations as a whole, some margin of discretion should be left open to allow the reasonable AWS commander to 'make a good faith judgment' regarding Attack Scope.[69] At the same time, the discussion in this section should provide a sufficient basis for superiors, legal analysts and post-hoc adjudicators to identify those situations where a commander clearly could *not* have considered that a particular choice of Attack Scope was reasonable.

## 4. APPLICATION TO TARGETING SCENARIOS

This section demonstrates, with the help of two scenarios, how the above methodology can be applied in practice to determine the scope of an AWS attack. This exercise also provides insights into how associated targeting obligations would be impacted.

### *How Many Attacks?*

To start, let us return to Scenario 1 and attempt to answer the question posed there: how many attacks is Commander-A responsible for today? For the 1200 sortie, the tanks are co-located, and the AWS is presumably poised to engage all objectives in short succession. For this situation, it is argued that Commander-A is entitled to analyse the sortie as a single Group-attack.[70] Commander-A only needs to verify the tanks' military nature once before the engagement, and may aggregate military advantage and collateral damage, etc. This is the same way in which an artillery volley against the same group of tanks would be characterised and assessed.

For the 1500 sortie, however, the tanks are geographically separated, presumably by a significant distance. In addition, the scenario makes clear that Commander-A is only sending *one* AWS to attack all four tanks, which would entail travel time that must be taken into consideration. It is argued that this operation cannot be treatedas a single Group-attack, but rather, that it must be treated as four Single-attacks. This entails separate proportionality calculations, separate attempts at mitigating collateral damage, etc. Crucially, Commander-A must continually monitor the *other* tanks while their AWS is occupied with the first, and cancel the subsequent strikes ifcircumstances legally demand it (e.g., if one tank rolls up to a market stall).[71] This is the same assessment that would need to be made were artillery to be used on the four

68 WB Huffman, 'Margin of Error: Potential Pitfalls of the Ruling in The Prosecutor v. Ante Gotovina' (2012) 211 Military Law Review 1, 17.

69 ibid 49.

70 Recall that for Group-attacks, the objectives must all belong to a Specific Target Group, which is the case for a tank platoon.

71 API (n 9) art 57(2)(b).

tanks; that is, they would constitute separate attacks and, thus, require distinct legal analyses.

This example shows how significantly Attack Scope impacts a commander's options for allowing AWS to operate autonomously. As hypothesised in Section 1, there is no simple answer as to whether autonomy is or is not allowed under IHL: it is always a question of whether said autonomy allowed is sufficiently narrow in terms of Attack Scope.

The answer to the question of how many attacks Commander-A is responsible for today is, thus, five: one Group-attack at 1200 hours and four Single-attacks at 1500.

We have now seen examples of Single- and Group-attacks. Can attacks also be of other tiers? A *Shot-attack* is a special Single-attack, where the AWS is only expected to make one engagement decision (e.g., 'kamikaze'-munitions[72]). Are offensive[73] *Signature-attacks* legally possible? While rarer, theoretically, yes—if the temporal and spatial conditions remain satisfied. Suppose a mix of enemy tanks and fuel trucks are moving through a small area, and the commander releases a swarm of AWS that can accurately identify and destroy both types of vehicles and rapidly destroy them. This would constitute a lawful Signature-level attack, but only if the fleet is sufficiently numerous to neutralise all objectives swiftly,[74] narrow geographical restrictions are applied, and there is little risk of changes in the environment during this period.

### *Assessing Proportionality*

For a second demonstration that focuses more on proportionality and the duty to cancel/suspend, consider the following:

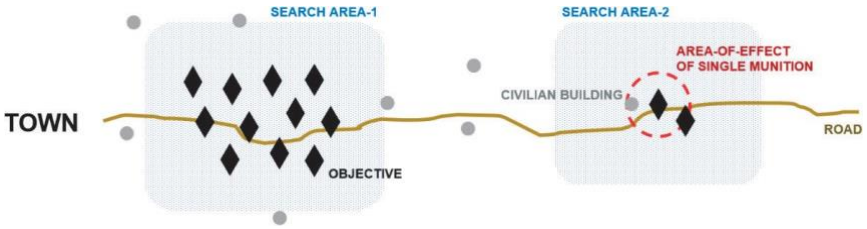
> **Scenario 2.** Enemy armour has been positively identified on a road as depicted in Figure 2. Commander-B considers sending AWS that can very reliably search and strike armour until they are neutralised, even when they take evasive action.

72 Note that if *multiple* munitions are sent (e.g. a kamikaze-swarm), the attack could potentially be qualified as a Group- or Signature-attack instead of individual Shot-attacks, depending on spatial and temporal proximity.

73 It was noted above that many defensive systems such as C-RAM and air defence fall in the Signature-tier, and are being used without controversy.

74 If only a single AWS is sent, it is likely too 'slow' since it must engage each objective in succession, and other objectives will move away, foreseeably decreasing spatial proximity.

**FIGURE 2:** ATTACK ON ROAD




There are no concerns regarding API Articles 51(4) and 57(2)(a)(i);[75] however, under current conditions, the building in Area-2 is likely to be damaged. Is this attack proportional? That depends on how we characterise this scenario in terms of Attack Scope, which admits two[76] possibilities:

(a) Two Group-attacks, comprising Area-1 and Area-2, respectively. For Area-2, one building would be damaged just to destroy two tanks, which is probably excessive.
(b) One Signature-attack, with the length of the road as the AWS's search area. One building would be damaged to destroy thirteen tanks. This is likely justifiable.

In such situations, commanders must carefully consider what Attack Scope to apply, as it significantly affects what actions are lawful under the proportionality rule. In the current scenario, whether Commander-B may count this as a single Signature-attack will likely depend on the particular circumstances. The map lacks a scale, but amalgamating the military advantage of Area-1 and Area-2 would be more justifiable if the distance between areas were 50 metres, compared to 500 metres. One can also consider the size of the fleet. If Commander-B releases multiple AWS simultaneously against Area-1 and Area-2, it would be easier to construe the scenario as a unified attack, compared to if a single AWS must 'finish' with Area-2 first before moving to Area-1. Once again, the temporal and spatial dimensions are decisive.

As theorised in Section 2, commanders will be tempted to define AWS attacks as broadly as possible because doing so carries significant practical benefits. In this example, if construed as two Group-attacks, the commander might have to delay or cancel the attack on Area-2, which may frustrate attempts to clear the road. Note that if clearing the road is *imperatively* important, then even a Group-attack on Area-2 could become justifiable, since the military advantage would derive not only from the destruction of matériel but also from the broader context (e.g., reinforcing the

75 The system is reliable enough to not be indiscriminate, and the objectives were clearly identified and validated by Commander-B.
76 Given the tanks clearly constitute two Specific Target Groups, we disregard the option of 13 Single-attacks.

town garrison that is about to collapse). Ultimately, this illustrates how judgments concerning Attack Scope (and the proportionality assessments that flow from them) are very context-dependent: commanders must utilise the guidelines presented above in good faith and not to justify *a priori* disproportionate attacks.

## CONCLUSION

Innovations in technology constantly require us to revisit existing IHL concepts. The notion of *attack* may initially seem unproblematic with respect to AWS, but closer inspection indicates that much legal indeterminacy exists with regard to how broadly commanders can define the *scope* of AWS attacks. It was hypothesised that uncertainty on the appropriate scope of AWS attacks is a major explanatory factor of the international disagreement on the level of autonomy that AWS are legally allowed to exhibit, and this paper is one attempt to re-frame the problem through a new lens that may advance the overall debate.

To address the problem of indeterminacy, this paper proposed a scaling methodology in the form of an Attack Scope spectrum, which can help commanders to determine in more transparent terms how they characterise AWS attacks in terms of scope. Two scenarios[77] were subsequently presented and discussed to demonstrate the methodology in practice. From this analysis, it was found that an attack can technically fall in any tier within this spectrum depending on the particular circumstances of the operation, but that the appropriate Attack Scope will depend on a balance of military and humanitarian interests. The military perspective dictates that Attack Scope must not be restricted too narrowly if this would render the execution of reasonable military operations and the application of the proportionality rule unworkable. The contrasting humanitarian perspective dictates that AWS attacks must be restricted in time and space so as not to jeopardise the protections that IHL grants to the civilian population through the principles of discrimination and proportionality. Ultimately, the guidelines presented in this paper sacrifice neither perspective – an approach that reflects the core balancing philosophy of IHL.[78]

Greater clarity and transparency concerning Attack Scope are liable to positively affect both the belligerent and the civilian. The AWS commander is strongly encouraged to consciously consider Attack Scope before launching any AWS attack. Knowing at which intervals they are imperatively required to perform legal assessments during their AWS's operational cycle removes legal ambiguity. Additionally, in cases where

77 The two scenarios presented in this work were designed to illustrate the concepts introduced in this paper in practical conditions. Future work may involve the study of more variations of toy scenarios to theorise, for example, if there are generally accepted limits in space and time which most reasonable commanders would consider as absolute boundaries with regard to stretching Attack Scope, which may further refine the general guidelines offered in Section 3.

78 AP Commentary (n 18) para 1389.

they are later called to account for any (incidental) harm caused by their AWS,[79] such reflection will allow them to justify how they delineated the contours of their attack(s) and explain for each attack why they did not deem the collateral damage to be excessive. For the civilian, clear boundaries to the extent to which commanders may extend the notion of attack positively impact their protection under IHL, in particular by requiring commanders to execute (re)validation and (re)assessment of proportionality at appropriate moments and by clarifying when commanders must maintain oversight over the *next* attack when AWS are used for successive engagements.

79 US Department of Defense (n 10) sct 5.10.2.2.