

ARTICLE

# Autonomous Weapons and the Right to Self-Defence

Agata Kleczkowska

Institute of Law Studies, Polish Academy of Sciences, Warsaw, Poland  
Email: [agata.kleczkowska@inp.pan.pl](mailto:agata.kleczkowska@inp.pan.pl)

## Abstract

This article focuses on the application of autonomous weapons (AWs) in defensive systems and, consequently, assesses the conditions of the legality of employing such weapons from the perspective of the right to self-defence. How far may humans exert control over AWs? Are there any legal constraints in using AWs for the purpose of self-defence? How does their use fit into the traditional criteria of self-defence? The article claims that there are no legal grounds to exclude AWs in advance from being employed to exercise the right to self-defence. In general, the legality of their use depends on how they were pre-programmed by humans and whether they were activated under proper circumstances. The article is divided into three parts. The first discusses how human control over AWs affects the legality of their use. Secondly, the article analyses the criteria of necessity and proportionality during the exercise of the right to self-defence in the context of the employment of AWs. Finally, the use of AWs for anticipatory, pre-emptive or preventive self-defence is investigated.

**Keywords:** anticipatory self-defence; autonomous weapons; self-defence; use of force

## 1. Introduction

It is too early to speak about ‘the arms race’ between states in the context of autonomous weapons (AWs). However, the largest arms-producing states (United States, United Kingdom, France, Russia, Italy, Japan, Israel, Germany, South Korea, India and China) already assume that ‘robotic technologies will contribute to shaping the future of warfare and that the acquisition of such technologies, particularly unmanned aerial systems (UASs), should be a

© The Author(s), 2023. Published by Cambridge University Press in association with the Faculty of Law, the Hebrew University of Jerusalem. This is an Open Access article, distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives licence (<http://creativecommons.org/licenses/by-nc-nd/4.0>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided that no alterations are made and the original article is properly cited. The written permission of Cambridge University Press must be obtained prior to any commercial use and/or adaptation of the article.

priority'.<sup>1</sup> The pioneer here is undoubtedly the United States: it started to work on the employment of artificial intelligence on the battlefield back in the 1980s. It has since gone a long way in the development of sophisticated technologies, including AWs, in its strategic policies. In 2016, the US Deputy Secretary of Defense announced that the 'technological sauce' of the Third Offset Strategy 'is going to advance in Artificial Intelligence (AI) and autonomy'.<sup>2</sup>

Autonomous weapons are considered to have especially high value when used in missions that require speed in decision making, analysis of heterogeneity and volumes of data, when the mission is dangerous, or when persistence and endurance are required.<sup>3</sup> Whether these functions will be used for defensive or offensive purposes depends on the tasks for which the weapon is used, the nature of their application and their reliability.<sup>4</sup> The aim of this article is to focus on the application of AWs in defensive systems and, consequently, to assess the conditions of the legality of their employment from the perspective of the right to self-defence. How far may humans exert control over such AWs? Are there any legal constraints in using such weapons for the purpose of self-defence? How does their use fit into the traditional criteria of self-defence? It is the aim of this article to answer these questions.

Before moving on to the merits, it is also important to define the scope of the article. In general, it may be said that AWs 'select and apply force to targets without human intervention'.<sup>5</sup> That definition is broad enough to allow 'consideration of the full range of relevant weapon systems, including existing weapons with autonomy in their critical functions that do not necessarily raise legal issues'.<sup>6</sup> Thus, this description includes (i) supervised AWs, using which humans oversee all the tasks performed by machines, (ii) semi-autonomous weapons, using which 'the machine performs a task and then stops and waits for approval from the human operator before continuing', in addition to (iii) fully autonomous weapons, using which a human operator

<sup>1</sup> Vincent Boulanin and Maaike Verbruggen, *Mapping the Development of Autonomy in Weapon Systems* (SIPRI 2017) 58.

<sup>2</sup> US Department of Defense, Deputy Secretary of Defense Bob Work, 'Remarks by Deputy Secretary Work on Third Offset Strategy', speech delivered in Brussels (Belgium), 28 April 2016, <https://www.defense.gov/News/Speeches/Speech/Article/753482/remarks-by-deputy-secretary-work-on-third-offset-strategy>.

<sup>3</sup> Department of Defense Science Board, 'Summer Study on Autonomy', June 2016, 12, <https://www.hsdl.org/?view&did=794641>.

<sup>4</sup> Zachary Kallenborn, 'Swords and Shields: Autonomy, AI, and the Offense-Defense Balance', *Georgetown Journal of International Affairs*, 22 November 2021, <https://gji.georgetown.edu/2021/11/22/swords-and-shields-autonomy-ai-and-the-offense-defense-balance>.

<sup>5</sup> International Committee of the Red Cross (ICRC), Position and Background Paper, 'ICRC Position on Autonomous Weapon Systems', 12 May 2021, 2, [https://www.icrc.org/en/download/file/166330/icrc\\_position\\_on\\_aws\\_and\\_background\\_paper.pdf](https://www.icrc.org/en/download/file/166330/icrc_position_on_aws_and_background_paper.pdf).

<sup>6</sup> 'Summary Report prepared by the International Committee of the Red Cross' in *Expert Meeting on Autonomous Weapon Systems: Implications of Increasing Autonomy in the Critical Functions of Weapons*, 15–16 March 2016 (ICRC 2016) 7, 8 (ICRC Summary Report).

pre-programmes the machine but does not choose specific targets.<sup>7</sup> Once the machine is activated, it will detect and engage targets on its own on the battlefield. Human operators are not able to supervise the operation and correct the action of the machine quickly enough in the event of the failure of the system. That is why this type of control is also known as ‘human out of the loop’.<sup>8</sup>

The role of humans in the targeting process may influence the legal acceptability of a particular weapons system;<sup>9</sup> this is why the aim of this article is to focus on this kind of AW where the use of the weapon has the lowest degree of human involvement. As a result of ever-evolving technology, it is difficult to create a comprehensive definition of an AW; it is rather understood that ‘[p]urely technical characteristics such as physical performance, endurance or sophistication in targeting acquisition and engagement’ are not sufficient to label a weapon as fully autonomous.<sup>10</sup> ‘Long-term developments regarding autonomous weapons are largely dependent on advances in the field of artificial intelligence’.<sup>11</sup> There are three types of artificial intelligence: (i) limited artificial intelligence, the one most widely used today; (ii) artificial general intelligence, ‘supposedly on a par with human intelligence’;<sup>12</sup> and (iii) artificial superintelligence, which ‘surpasses human intelligence’.<sup>13</sup> Even if some militaries and experts express doubts about whether fully autonomous weapons which use artificial general intelligence and artificial superintelligence already exist today,<sup>14</sup> they agree that the era of such weapons is quickly approaching.<sup>15</sup> Conversely, fully autonomous weapons with limited artificial intelligence are currently broadly used – for instance, in systems used for air defence.

<sup>7</sup> Paul Scharre, ‘Autonomous Weapons and Operational Risk’, *Ethical Autonomy Project*, February 2016, 9, <https://www.cnas.org/publications/reports/autonomous-weapons-and-operational-risk>.

<sup>8</sup> *ibid* 9–10, 20.

<sup>9</sup> ICRC Summary Report (n 6) 9.

<sup>10</sup> ‘Report of the 2018 Session of the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems’, 23 October 2018, UN Doc C/W/GGE.1/2018/3, para 22b.

<sup>11</sup> Advisory Council on International Affairs (AIV) and Advisory Committee on Issues of Public International Law (CAVV), ‘Autonomous Weapon Systems: The Need for Meaningful Human Control’, October 2015, No 97 AIV/No 26 CAVV, 16 (AIV and CAVV).

<sup>12</sup> *ibid*.

<sup>13</sup> *ibid* 17.

<sup>14</sup> Gerrit De Vynck, ‘The U.S. Says Humans Will Always Be in Control of AI Weapons. But the Age of Autonomous War Is Already Here’, *The Washington Post*, 7 July 2021, <https://www.washingtonpost.com/technology/2021/07/07/ai-weapons-us-military>.

<sup>15</sup> eg, the Australian Jaeger-C uncrewed combat vehicle: David Hambling, ‘Australian Army Getting Bulletproof Swarming Attack Robots’, *Forbes*, 4 November 2021, <https://www.forbes.com/sites/davidhambling/2021/11/04/australian-army-gets-bulletproof-attack-robots/?sh=174d19fe463b>; Turkish KARGU-2 drones: STM, ‘KARGU: Combat Proven Rotary Wing Loitering Munition System’, <https://www.stm.com.tr/en/kargu-autonomous-tactical-multi-rotor-attack-uav>; or the US MQ-9 Reaper equipped with Agile Condor system: ‘Agile Condor High-Performance Embedded Computing’, <https://www.srcinc.com/pdf/intelligence-surveillance-reconnaissance-agile-condor>.pdf; General Atomics, ‘GA-ASI Awarded Smart Sensor Contract’, 24 November 2020, <https://www.ga.com/ga-asi-awarded-smart-sensor-contract>. See also Kyle Hiebert, ‘Are Lethal Autonomous Weapons Inevitable? It Appears So’, 27 January 2022, <https://www.cigionline.org/articles/are-lethal-autonomous-weapons-inevitable-it-appears-so>.

They include ground- and ship-based installations against missiles, rockets, artillery, mortars, aircraft, unmanned systems and high-speed boats. Examples include the Dutch Goalkeeper Close-in Weapon System, the Israeli Iron Dome, the Russian Kashtan Close-in Weapon System, the German Nächstbereichsschutzsystem MANTIS, the US Phalanx Close-in Weapon System, and the Chinese Type 730 and Type 1130 Close-in Weapon Systems.<sup>16</sup> Such understanding of fully autonomous weapons will be adopted in the latter part of the article.<sup>17</sup>

The article claims that there are no legal grounds to exclude AWs in advance from being employed to exercise the right to self-defence. In general, the legality of the use of AWs depends on how they were pre-programmed by humans and whether they were activated under proper circumstances. Because designers of AWs are not able to envisage all conditions under which the weapons will work, it is possible that the use of AWs will result in breaches of the criteria of necessity and proportionality. Moreover, the vagueness of the concept of anticipatory self-defence and the unpredictability of AWs in action mean that this type of weapon should not be employed for the purposes of the anticipatory strike, for the reasons that it is likely to lead to a violation of international law.

This article is divided into three parts. The first (Section 2) discusses how human control over AWs affects the legality of their use. Secondly, the article analyses (in Section 3) the criteria of necessity and proportionality during the exercise of the right to self-defence in the context of the employment of AWs. Finally, the use of AWs for anticipatory, pre-emptive, or preventive self-defence is investigated (Section 4).

## 2. The impact of human control over the development, deployment and operation of AWs

The acceptability of the legality of the use of AWs depends, mostly, on the degree of human control that is exercised over them in the course of three different stages: the development, deployment, and operational phases.<sup>18</sup> It is the aim of this section to analyse these phases from the perspective of the human and machine interactions that occur during each of these steps and relate them to requirements set by *jus ad bellum*.

### 2.1. Human control during the development phase

During the first phase, humans carry out 'an extensive planning ... in which they set the overall goals, gather intelligence, select and develop targets, identify the most suitable weapon, and decide in what circumstances and under

<sup>16</sup> Neil Davison and Gilles Giacca, 'Background Paper prepared by the International Committee of the Red Cross March 2016' in ICRC *Expert Meeting* (n 6) 72, 72–73.

<sup>17</sup> Whenever the article uses the term 'autonomous weapon' from this point, it should be understood as relating to fully autonomous weapons.

<sup>18</sup> The scope of human control over AWs is also critical when it comes to acceptance of their use by public opinion: AIV and CAVV (n 11) 32.

what preconditions to employ a particular weapon'.<sup>19</sup> In short, this phase consists of pre-programming the AW.

It seems that it is this phase, rather than the operational phase, that is critical for the legality of the employment of AWs.<sup>20</sup> Against some myths spread by opponents of AWs, the machine does not take decisions discretely; its actions are the consequence of the decisions taken by humans in the development phase.<sup>21</sup> For instance, it is sometimes submitted that AWs 'select' their targets, while it would be more precise to state that they only 'detect' them. They are equipped with a target-identification library and various sensor capabilities. Only when their sensors recognise the target as matching the predefined criteria may they fire upon it.<sup>22</sup> In the same vein, AWs do not 'apply legal norms' but 'they are simply implementing a process that human commanders anticipate in their assessment of the legality of a planned attack'.<sup>23</sup> What is obviously non-programmable 'is conscience, common sense, intuition, and other essential human qualities'.<sup>24</sup> However, some claim that AWs may be much stricter in complying with international law than humans. This is because they are deprived of emotions and other mental and cognitive constraints characteristic of people.<sup>25</sup>

Thus, during the development process, humans should ensure that the machine will operate in compliance with both *jus ad bellum* and *jus in bello* requirements. Subsequently, an AW should be tested in a realistic environment in order to ascertain how the machine acts in varying circumstances and the risks that are connected with its deployment.<sup>26</sup> The conclusions obtained during the tests will be of utmost importance for the human operators of the AW in the next phase.

## 2.2. Human control during the deployment phase

In the second phase, the human operator decides on the activation of the AW. This phase is burdened with particular responsibility as, after activation, a human will no longer be able to intervene in the functioning of the AW.

<sup>19</sup> Merel Ekelhof, 'Human Control in the Targeting Process: Speaker's Summary' in *ICRC Expert Meeting* (n 6) 53, 55.

<sup>20</sup> AIV and CAVV (n 11) 10.

<sup>21</sup> See also statement made by Israel in Campaign to Stop Killer Robots, 'Report on Activities: Convention on Conventional Weapons Second Informal Meeting of Experts on Lethal Autonomous Weapons Systems', 13–17 April 2015, 13, [https://www.stopkillerrobots.org/wp-content/uploads/2013/03/KRC\\_CCWx2015\\_Report\\_4June2015\\_uploaded.pdf](https://www.stopkillerrobots.org/wp-content/uploads/2013/03/KRC_CCWx2015_Report_4June2015_uploaded.pdf).

<sup>22</sup> Heather Roff, 'Sensor-Fused Munitions, Missiles and Loitering Munitions: Speaker's Summary' in *ICRC Expert Meeting* (n 6) 33, 33–34.

<sup>23</sup> Richard Moyes, 'Meaningful Human Control over Individual Attacks: Speaker's Summary' in *ICRC Expert Meeting* (n 6) 46, 47.

<sup>24</sup> Mary Ellen O'Connell, 'Banning Autonomous Killing: The Legal and Ethical Requirement that Humans Make Near-Time Lethal Decisions' in Matthew Evangelista and Henry Shue (eds), *The American Way of Bombing: Changing Ethical and Legal Norms from Flying Fortresses to Drones* (Cornell University Press 2014) 224, 232.

<sup>25</sup> Rosa Brooks, 'In Defense of Killer Robots', *Foreign Policy*, 18 May 2015, <https://foreignpolicy.com/2015/05/18/in-defense-of-killer-robots>.

<sup>26</sup> AIV and CAVV (n 11) 36.

Thus, the time, location and circumstances of the activation of the weapon may be crucial for both combatants and the civilian population. For this reason, any activation of an AW should be preceded by a carefully planned scheme that allows for the avoidance of collateral damage.<sup>27</sup>

The human operator who activates an AW needs to be 'well trained and have a good understanding of the system and how it interacts with the environment in which it is deployed'.<sup>28</sup> This operator should be aware of the technical and other limitations of a particular AW. Moreover, the operator needs to be able 'to foresee what might happen without being able to predict precisely what will happen'.<sup>29</sup> Obviously, the predictability of the given weapon also performs a huge role as it allows the human operator to assess the advantages and disadvantages of the deployment of an AW in the given environment. A combination of training, an understanding of the predictability of the weapon, in addition to the analysis of all available information, should allow a human operator to decide whether a specific case of the use of an AW would not violate international law.<sup>30</sup>

### 2.3. Human control in the operational phase

It is predominantly the third phase that decides about the full autonomy of a weapon: that is, a weapon is fully autonomous when there is minimal or no direct human control at this stage. It means that humans have pre-programmed the weapon and decided when to launch it but the weapons system autonomously decides every time whether to use force against a specific target.<sup>31</sup>

Some experts and organisations promote the term 'meaningful human control'.<sup>32</sup> As explained by Tim McFarland and Jai Galliot:<sup>33</sup>

The idea is that some form of human control over the use of force is required, and that human control cannot be merely a token or a formality; human influence over acts of violence by a weapon system must be sufficient to ensure that those acts are done only in accordance with human designs and, implicitly, in accordance with legal and ethical constraints.

The 'meaningful human control' is of special significance for the operational phase, after the AW is activated. As the proponents of this concept claim, human control is indispensable additionally over the process of the use of

<sup>27</sup> *ibid* 9–10.

<sup>28</sup> *ibid* 12.

<sup>29</sup> *ibid* 33.

<sup>30</sup> *ibid* 34.

<sup>31</sup> ICRC Summary Report (n 6) 16.

<sup>32</sup> Tim McFarland and Jai Galliot, 'Understanding AI and Autonomy: Problematising the Meaningful Human Control Argument against Killer Robots' in Jai Galliot, Duncan MacIntosh and Jens David Ohlin (eds), *Lethal Autonomous Weapons: Re-examining the Law and Ethics of Robotic Warfare Understanding AI and Autonomy* (Oxford University Press 2021) 41, 41–50.

<sup>33</sup> *ibid* 45.

force by AWs, to ensure full compliance of machine actions with international law. However, it is hard to explain how the exercise of ‘meaningful human control’ could be helpful in cases where the violation of international law is the result of an intentional human act or a malfunction, or when the unavoidable collateral damage occurs. Importantly, none of these wrongs is unique for AWs, as the same kind of violation may occur when a human operator is using a fully controllable weapon.<sup>34</sup>

#### 2.4. Conclusions

The degree of human control over AWs at different stages of their usage determines both the scope of the autonomy of these weapons, as well as the acceptance of the legality of their deployment. Needless to say, human involvement in the development phase is indispensable. In general, the data pre-programmed into the AW at this stage and the decision to activate the weapon under given circumstances resolve the issue of whether the use of force by the machine will comply with international law. The action of the machine in the operational phase is purely the consequence of the human choices made in the two previous stages. Thus, it is not the involvement of the human in the operational stage that decides on the legality of the AW employment, and it can hardly mitigate the wrongful decisions made earlier.

### 3. The application of the criteria of necessity and proportionality in cases of the use of autonomous weapons

Although Article 51 of the UN Charter<sup>35</sup> does not mention that ‘[t]he submission of the exercise of the right of self-defence to the conditions of necessity and proportionality is a rule of customary international law’,<sup>36</sup> the two criteria are inseparably linked, partially overlap, and cannot be reviewed separately.<sup>37</sup> This is so for the reasons that, ultimately, they have a joint objective: ‘that defensive action should be constrained to the halting and repelling of an attack, and exceptionally ... to the preventing of their reoccurrence’.<sup>38</sup>

In general terms, one may say that the principle of necessity responds to the question whether a specific measure is necessary to achieve a legitimate purpose of self-defence, whereas the principle of proportionality answers

<sup>34</sup> *ibid* 51.

<sup>35</sup> Charter of the United Nations (entered into force 24 October 1945) 1 UNTS XVI.

<sup>36</sup> ICJ, *Legality of the Threat or Use of Nuclear Weapons*, Advisory Opinion [1996] ICJ Rep 226, [41]. See also ICJ, *Military and Paramilitary Activities in and against Nicaragua (Nicaragua v United States of America)* Merits, Judgment [1986] ICJ Rep 14, [176].

<sup>37</sup> Christine Gray, *International Law and the Use of Force* (4th edn, Oxford University Press 2018) 159; Albert Randelzhofer and Georg Nolte, ‘Article 51’ in Bruno Simma and others (eds), *The Charter of the United Nations: A Commentary* (Oxford University Press 2012) 1397, 1425; International Law Association, ‘Final Report on Aggression and the Use of Force’, Sydney Conference, 2018, 12.

<sup>38</sup> Tom Ruys, *‘Armed Attack’ and Article 51 of the UN Charter* (Cambridge University Press 2010) 123–24; see also International Law Association (n 37) 11.

the question how far a specific measure may go in order to achieve such a purpose.<sup>39</sup>

However, to examine how the criteria of necessity and proportionality are applicable to the case of AWs, in-depth analysis is indispensable.

### 3.1. Necessity and proportionality – general remarks

#### 3.1.1. Necessity

Starting from the characteristics of necessity, one has to observe that, to comply with this criterion, self-defence has to be the *ultima ratio*. To put it differently, there must be ‘no choice of means’;<sup>40</sup> a state must have no other means of halting an attack other than recourse to armed force.<sup>41</sup> Thus, this condition is not fulfilled and, consequently, the self-defence is illegal if a state could halt an attack by using peaceful measures such as an admonition, protest or diplomatic representation, could offer arbitration, or use the procedure of settling disputes provided by an international organisation to which it belongs.<sup>42</sup>

Obviously, one has to agree with Oscar Schachter when he writes that ‘[t]here is no legal rule that a state must turn the other cheek because of its obligation under Article 2 (3) to seek peaceful settlement’.<sup>43</sup> It may be the case that an armed response is the only viable option to repel an armed attack, and there is ‘no reasonable prospect of the efficacy of peaceful measures of settlement’.<sup>44</sup>

The second component is the requirement of ‘immediacy’,<sup>45</sup> which refers to the time when a state, responding to an attack, launches its action of deterrence. ‘[T]he longer the period between the armed attack and the response, the more pressure there will be on the State concerned to resolve the matter by peaceful means’.<sup>46</sup> Moreover, the more time that elapses between the armed attack and the response to it, the harder it will become to distinguish between self-defence and illegal reprisals if the state still decides to react militarily.<sup>47</sup>

Thirdly, there is the criterion of exclusive purpose: the measures undertaken must be of a genuinely defensive character, ‘that is, designed to put an end to an armed attack’.<sup>48</sup> One may bind this criterion with the requirement

<sup>39</sup> Randelzhofer and Nolte (n 37) 1425–26.

<sup>40</sup> Bin Cheng, *General Principles of Law as Applied by International Courts and Tribunals* (Grotius 1987) 95.

<sup>41</sup> International Law Commission, ‘Addendum to the Eighth Report on State Responsibility, by Mr. Roberto Ago’, UN Doc A/CN.4/318/ADD.5-7, para 120.

<sup>42</sup> Oscar Schachter, ‘International Law in Theory and Practice General Course in Public International Law’ (1985) 178 *Recueil de Cours* 1, 152–53.

<sup>43</sup> *ibid* 154. Schachter refers to Article 2(3) of the UN Charter (n 35), which states: ‘All Members shall settle their international disputes by peaceful means in such a manner that international peace and security, and justice, are not endangered’.

<sup>44</sup> Judith Gardam, *Necessity, Proportionality and the Use of Force by States* (Cambridge University Press 2004) 152–53.

<sup>45</sup> Ruys (n 38) 95.

<sup>46</sup> Gardam (n 44) 150–51.

<sup>47</sup> Ruys (n 38) 99.

<sup>48</sup> Olivier Corten, *The Law Against War* (Hart 2010) 484.

to choose an appropriate target to repel an armed attack.<sup>49</sup> States, when deciding on self-defence actions in smoke-filled rooms, sometimes choose to attack not the immediate source of an armed attack against the state but objects on the territory of the adversary that are especially sensitive or valuable to that state.<sup>50</sup> However, the use of force for a purpose other than that of a defensive response amounts to a violation of international law. For this reason, this criterion also needs to be carefully obeyed.

### 3.1.2. Proportionality

The function of proportionality 'is to serve as a constraint on the scale and effects of defensive action'.<sup>51</sup> However, one may approach the question of the scale and effects of self-defence in two ways. The first position holds that 'there must be some sort of equation between the gravity of the armed attack and the defensive response, in terms of relative casualties, damage caused and weapons used' (quantitative approach).<sup>52</sup> The other view is that 'proportionality must be evaluated by reference to the aim of the defensive action' (functional approach).<sup>53</sup> Tom Ruys suggests, however, that the most appropriate position would be that 'any defensive action must be reasonably proportionate in scale and nature to the armed attack that provoked it'.<sup>54</sup> Any deviation from this line of reasoning needs to be justified by the specific objective of the defensive action.<sup>55</sup> That is the approach that will be adopted in this article.

What should be taken into account in assessing proportionality? Firstly, it concerns the choice of methods of warfare.<sup>56</sup> Importantly, proportionality does not mean that the victim of an armed attack must respond with exactly the same weapon that the attacker used in order to satisfy this criterion. On the contrary, in theory there are no constraints preventing the victim from moving 'to a higher mode of weaponry and to a greater degree of firepower'.<sup>57</sup> Moreover, no type of weapon that the victim could use to defend itself is categorically excluded in advance.<sup>58</sup> The International Court of Justice (ICJ) ruled that the proportionality principle does not even exclude the use of nuclear

<sup>49</sup> Ruys (n 38) 108.

<sup>50</sup> In reaction to the arrest of SS *Mayaguez* by Cambodia in 1975, the US, inter alia, struck an airfield, damaging the runway and hangar, and destroying numerous aircraft, as well as targeted two naval facilities. President Ford justified the US actions with the right to self-defence, even though the Americans did not have any information about the fate of the vessel or crew: Joseph Eldred, 'The Use of Force in Hostage Rescue Missions' (2008) 56 *Naval Law Review* 251, 262.

<sup>51</sup> Ruys (n 38) 110.

<sup>52</sup> *ibid* 111.

<sup>53</sup> *ibid* 112.

<sup>54</sup> *ibid* 116–17.

<sup>55</sup> *ibid*.

<sup>56</sup> Gardam (n 44) 168.

<sup>57</sup> DP O'Connell, *The Influence of Law on Sea Power* (Manchester University Press 1975) 55. See also Randelzhofer and Nolte (n 37) 1426.

<sup>58</sup> Daniele Amoroso, 'Jus in Bello and Jus ad Bellum Arguments against Autonomy in Weapons Systems: A Re-appraisal' (2017) 43 *Questions in International Law: Zoom-in* 5, 29.

weapons in self-defence.<sup>59</sup> However, when choosing the measures of self-defence, a state must take into consideration the fact that the weapon it wishes to use cannot prompt a scale of destruction disproportionate to the attack, in addition to the fact that the state is not allowed to 'employ its more destructive weaponry if it can achieve the legitimate objectives of self-defence with the lesser weapons available to it'.<sup>60</sup>

Secondly, the choice of targets also influences the assessment of proportionality. A state, in using force in self-defence, must take into account the 'anticipated overall scale of civilian casualties, the level of destruction of enemy forces, and finally damage to the territory, the infrastructure of the target state and the environment generally'.<sup>61</sup> One example of a disproportionate response to an armed attack can be found in the *Oil Platforms* judgment. The ICJ observed that 'the destruction of two Iranian frigates and a number of other naval vessels and aircraft ... [a]s a response to the mining, by an unidentified agency, of a single United States warship, which was severely damaged but not sunk' cannot be regarded 'as a proportionate use of force in self-defence'.<sup>62</sup> Also, in *Military and Paramilitary Activities in and against Nicaragua* the ICJ found that the US activities 'relating to the mining of the Nicaraguan ports and the attacks on ports, oil installations, etc' were not a proportionate reaction to 'the aid received by the Salvadorian armed opposition from Nicaragua', even if the scale of that aid was uncertain.<sup>63</sup>

### 3.2. Necessity and proportionality in the context of deployment of autonomous weapons

The doctrine of law and military experts are not unanimous on whether AWs may satisfy the criteria of necessity and proportionality,<sup>64</sup> including proper evaluation of the involvement of humans, and distinguishing civilian and military objects when determining the target of the attack.<sup>65</sup> The fundamental question is whether an AW is able, to the same extent as a human being, to

<sup>59</sup> ICJ, *Legality of the Threat or Use of Nuclear Weapons*, Advisory Opinion [1996] ICJ Rep 226, [42].

<sup>60</sup> Christopher Greenwood, 'Self-Defence and the Conduct of International Armed Conflict' in Yoram Dinstein (ed), *International Law at a Time of Perplexity: Essays in Honour of Shabtai Rosenne* (Martinus Nijhoff 1989) 273, 280.

<sup>61</sup> Gardam (n 44) 168.

<sup>62</sup> ICJ, *Oil Platforms (Islamic Republic of Iran v United States of America)* Merits, Judgment [2003] ICJ Rep 161, [77].

<sup>63</sup> *Military and Paramilitary Activities in and against Nicaragua* (n 36) 237. See also ICJ, *Armed Activities on the Territory of the Congo (Democratic Republic of the Congo v Uganda)*, Judgment [2005] ICJ Rep 168, [147] ('The Court cannot fail to observe, however, that the taking of airports and towns many hundreds of kilometres from Uganda's border would not seem proportionate to the series of transborder attacks it claimed had given rise to the right of self-defence, nor to be necessary to that end').

<sup>64</sup> For views in support of the position that AWs may comply with these criteria see, eg, Donovan Phillips, 'The Automation of Authority: Discrepancies with *Jus ad Bellum* Principles' in Galliot, Macintosh and Ohlin (n 32) 159, 162.

<sup>65</sup> Such view was also presented in ICRC Summary Report (n 6) 9. The question of whether AWs used in self-defence are able to comply with rules of international humanitarian law is important – in the words of the ICJ, 'a use of force that is proportionate under the law of self-defence, must, in order to be lawful, also meet the requirements of the law applicable in armed conflict which

consider all the factors that contribute to the fulfilment of these requirements. It not only refers to the development phase – that is, it is not only a technical problem of pre-programming the criteria of necessity and proportionality into AWs – it is also a question of whether, even if equipped with all the necessary sensors and data, AWs would be able to assess a specific situation properly according to the criteria of necessity and proportionality.

In one respect the fact that AWs in the operational phase take decisions without human interference may have a number of advantages. AWs, being connected to systems that provide them with a significant number of different kinds of data, may collect and process information in a much shorter time than humans can; thus, when a decision needs to be taken within fractions of seconds, machines will respond more quickly.<sup>66</sup> Under this time pressure, AWs may also choose the targets more effectively. For example, ‘the algorithm might recommend that the defending state target a particular command centre rather than a particular missile installation’.<sup>67</sup> As mentioned above, the fact that AWs are deprived of human emotions may also be a huge asset. ‘A robot cannot hate, cannot fear, cannot be hungry or tired and has no survival instinct’.<sup>68</sup> Moreover, in contrast to humans, a robot cannot deliberately choose not to comply with international law. Thus, it is more probable that the AW would abide by legal rules, including humanitarian principles.<sup>69</sup>

Conversely, however, many experts claim that AWs are unable to properly assess the criteria of necessity and proportionality; these objections seem to come at two levels. Firstly, it is doubtful whether it is possible to pre-programme AWs so they could comply with these conditions. To achieve that, it would mean that the designers of such a weapon would have to predict every single scenario in which the machine may be deployed. It is important to be aware of the fact that the fulfilment of the criteria of ‘proportionality [and necessity] of any particular attack depend ... on conditions at the time of the attack, and not at the moment of design or deployment of a weapon’.<sup>70</sup> Secondly, AWs are not only deprived of human emotions but also of other abilities that may turn out to be crucial for the proportionality and necessity of an attack. The algorithms that are prepared for the AW are not able to accommodate information inherent to the life and professional experience of a military commander. AWs may provide perfect calculations based on objective information and statistical data; the results, however, may be different from human

---

comprise in particular the principles and rules of humanitarian law’: *Legality of the Threat or Use of Nuclear Weapons* (n 59) [42].

<sup>66</sup> Martin Hagström, ‘Characteristics of Autonomous Weapon Systems: Speaker’s Summary’ in *ICRC Expert Meeting* (n 6) 23, 24.

<sup>67</sup> Ashley Deeks, Noam Lubell and Daragh Murray, ‘Machine Learning, Artificial Intelligence, and the Use of Force by States’ (2019) 10 *Journal of National Security Law and Policy* 1, 10.

<sup>68</sup> Marco Sassóli, ‘Autonomous Weapons and International Humanitarian Law: Advantages, Open Technical Questions and Legal Issues To Be Clarified’ (2014) 90 *International Law Studies* 308, 310.

<sup>69</sup> Brooks (n 25).

<sup>70</sup> Bonnie Docherty and others, ‘Making the Case: The Dangers of Killer Robots and the Need for a Preemptive Ban’, *Human Rights Watch*, 9 December 2016, 7, [https://www.hrw.org/report/2016/12/09/making-case/dangers-killer-robots-and-need-preemptive-ban#\\_ftn12](https://www.hrw.org/report/2016/12/09/making-case/dangers-killer-robots-and-need-preemptive-ban#_ftn12).

judgement of the situation.<sup>71</sup> Machines will not use the life-long experience or intuition that human commanders have;<sup>72</sup> they do not have 'clear understanding of personal and organizational values',<sup>73</sup> and are not capable of subjecting interpreted information to thorough reflection.<sup>74</sup>

When it comes specifically to proportionality, it is doubtful whether an AW could balance different values,<sup>75</sup> such as human life and the security of the state. The AW would have to adjust to the threat of the scale of force used. The question, however, is whether it can aptly recognize a threat; what if the nervous conduct of a frightened civilian is interpreted by the AW as a potential threat?<sup>76</sup>

To address necessity, AWs would have to be 'able to judge the degree to which life was at risk, and be able to select and implement alternatives to lethal force, such as negotiation and capture'.<sup>77</sup> Moreover, some experts claim that the deployment of AWs could contravene some fundamental human rights, including the right to life. Additionally, this is also because AWs lack empathy and compassion, which are important human inhibitors to the taking of human lives.<sup>78</sup>

### 3.3. Conclusions

The criteria of necessity and proportionality determine the legality of the use of force in self-defence. To fulfil these criteria, it is necessary not only to respond with similar force and strength. Proper application of necessity and proportionality requires the examination of many factors and not just simple, quantitative analysis. While it is true that AWs may analyse a huge amount of information within a short time and take decisions more effectively than humans, properly fulfilling the criteria of necessity and proportionality may, on occasion, also require human judgement of the situation, composed of experience, emotions and morality, something that AWs will possibly never be able to acquire.

## 4. Deployment of autonomous weapons for anticipatory, pre-emptive or preventive self-defence

Article 51 of the UN Charter states: 'Nothing in the present Charter shall impair the inherent right of individual or collective self-defence if an armed

<sup>71</sup> See also Bonnie Docherty, 'The Need for and Elements of a New Treaty on Fully Autonomous Weapons', *Human Rights Watch*, 1 June 2020, <https://www.hrw.org/news/2020/06/01/need-and-elements-new-treaty-fully-autonomous-weapons>.

<sup>72</sup> ICRC Summary Report (n 6) 20.

<sup>73</sup> John Stroud-Turp, 'Lethal Autonomous Weapon Systems (LAWS): Speaker's Summary' in ICRC *Expert Meeting* (n 6) 57, 58.

<sup>74</sup> Bonnie Docherty and others, 'Shaking the Foundations: The Human Rights Implications of Killer Robots', *Human Rights Watch*, 12 May 2014, 12, <https://www.hrw.org/report/2014/05/12/shaking-foundations/human-rights-implications-killer-robots>.

<sup>75</sup> ICRC Summary Report (n 6) 16.

<sup>76</sup> Docherty and others (n 74) 10–11.

<sup>77</sup> Nathalie Weizmann, 'Autonomous Weapon Systems under International Law' (2014) 8 *Academy Briefing* 1, 12.

<sup>78</sup> Docherty and others (n 74) 2, 6.

attack occurs against a Member of the United Nations'. In line with the explicit reading of Article 51, a number of authors claim that self-defence is legal only after an armed attack has taken place.<sup>79</sup> To refer this scenario to the deployment of AWs would mean that the weapons should be programmed in such a way as to make a decision about a strike only after the state is attacked.

However, apart from this straightforward scenario, other options are possible. One variable covers situations when an armed attack is considered to have started, even if the material (kinetic) effects of the attack have not yet occurred in the territory of the attacked state. An example may be a situation when a ballistic missile, launched by one state against the other, has not yet reached the target. An armed attack has nevertheless started or, in the words of Yoram Dinstein, 'once a button is pressed, or a trigger is pulled, the act is complete (while impact is a mere technicality)'.<sup>80</sup> Thus, the attacked state is allowed to resort to interceptive self-defence.

The application of AWs in self-defence in such cases does not raise substantial doubts. A number of AWs are designed specifically to intercept missiles. This is the case with the Iron Dome in Israel, which is 'a type of counter-rocket, artillery and mortar weapon system capable of intercepting multiple targets at short range'.<sup>81</sup> It is also the function of Terminal High-Altitude Area Defence, which 'detects and tracks an incoming missile, calculates its trajectory and then attacks it with an interceptor missile'.<sup>82</sup> Similar functions are performed by the SeaRAM Close-in Weapon System, a ship-based system. Other, autonomous systems which can detect incoming missiles and fire blast, munition or interceptor missiles include the German Advanced Modular Armour Protection Active Defence System, Russian Arena, South African LEDS-150, the United States' Quick Kill, Israeli Trophy, and Ukrainian Zaslouk.<sup>83</sup> Because all these weapons react to the ongoing armed attack against the state and are not designed to strike first, their legality is indisputable.

Despite the explicit reading of the UN Charter, several scholars suggest that self-defence is legal not only when an armed attack has occurred, but that international law also allows the use of force in self-defence when there is only the threat of an armed attack.<sup>84</sup> Self-defence before an armed attack

<sup>79</sup> Authors who also claim that self-defence is legal only after an armed attack occurs are, eg, Ruys (n 38) 57–60; Corten (n 48) 407; James Crawford, *Brownlie's Principles of Public International Law* (8th edn, Oxford University Press 2012) 750–51; Antonio Cassese, 'Article 51' in Jean-Pierre Cot and Alain Pellet (eds), *La Charte des Nations Unies* (Economica 1991) 775, 777–80; Hilaire McCoubrey and Nigel D White, *International Law and Armed Conflict* (Dartmouth 1992) 90.

<sup>80</sup> Yoram Dinstein, *War, Aggression and Self-Defence* (4th edn, Cambridge University Press 2005) 190.

<sup>81</sup> ICRC Summary Report (n 6) 10.

<sup>82</sup> *ibid.*

<sup>83</sup> Davison and Giacca (n 16) 72–73.

<sup>84</sup> eg, DW Bowett, 'Collective Self-Defence under the Charter of the United Nations' (1955–1956) 32 *British Yearbook of International Law* 67, 148; Marco Roscini, 'Threats of Armed Force and Contemporary International Law' (2007) 54 *Netherlands International Law Review* 229, 272; Louis Henkin, *How Nations Behave: Law and Foreign Policy* (Columbia University Press 1979) 140–44; Ian Brownlie, *International Law and the Use of Force by States* (Oxford University Press 1963) 278–29; Jaroslav Žourek, *L'Interdiction de l'Emploi de la Force en Droit International* (AW Sijthoff 1974) 101.

takes place may be understood in very different ways; this type is mirrored by terms used to describe such self-defence.

The notion of anticipatory self-defence is usually used to discuss 'a use of force against an imminent threat of armed attack'.<sup>85</sup> After Daniel Webster's formula, one can repeat that imminence takes place when the threat is 'instant, overwhelming, leaving no choice of means, and no moment for deliberation'.<sup>86</sup> However, states tend to interpret 'imminence' in varying ways, including when an armed attack is not 'about to happen very soon'.<sup>87</sup> For this reason, it is currently postulated that imminence 'cannot be construed by reference to a temporal criterion only, but must reflect the wider circumstances of the threat'.<sup>88</sup> Those who wish to restrict the possible abuses of the right to self-defence but who, concurrently, agree that states are not forced to wait until they are attacked in order to defend themselves, link the criterion of imminence with the requirement of necessity 'inasmuch as it requires that there be no time to pursue non-forcible measures with a reasonable chance of averting or stopping the attack'.<sup>89</sup>

Conversely, supporters of the right to pre-emptive or preventive self-defence claim that states have the 'right to respond to threats which have not yet crystallized but which might materialise at some time in the future'.<sup>90</sup> It is often claimed that the most striking example of this approach is the so-called Bush doctrine – views on the imminence of threats and legality of the use of force presented in the US National Security Strategy of 2002. The US declared, *inter alia*, that:<sup>91</sup>

[we] can no longer solely rely on a reactive posture as we have in the past. The inability to deter a potential attacker, the immediacy of today's threats, and the magnitude of potential harm that could be caused by our adversaries' choice of weapons, do not permit that option. We cannot let our enemies strike first. ... Legal scholars and international jurists often conditioned the legitimacy of preemption on the existence of an imminent threat—most often a visible mobilization of armies, navies, and air forces preparing to attack. We must adapt the concept of imminent threat to the capabilities and objectives of today's adversaries. ...

<sup>85</sup> Alan L Schuller, 'Inimical Inceptions of Imminence: A New Approach to Anticipatory Self-Defense under the Law of Armed Conflict' (2014) 18 *UCLA Journal of International Law and Foreign Affairs* 161, 168.

<sup>86</sup> 'The United States Secretary of State Daniel Webster, note dated 24 Apr. 1841' (1841) 29 *British and Foreign State Papers* 1137, 1137–38.

<sup>87</sup> Christian Henderson, *The Use of Force and International Law* (Cambridge University Press 2018) 297.

<sup>88</sup> Elizabeth Wilmschurst, 'Principles of International Law on the Use of Force by States in Self-Defence', Chatham House, October 2005, 8, <https://www.chathamhouse.org/sites/default/files/publications/research/2005-10-01-use-force-states-self-defence-wilmschurst.pdf>.

<sup>89</sup> *ibid* 7.

<sup>90</sup> *ibid* 9.

<sup>91</sup> US President George W Bush, 'The National Security Strategy', September 2002, Ch V, <https://georgewbush-whitehouse.archives.gov/nsc/nss/2002/nss5.html>.

The United States has long maintained the option of preemptive actions to counter a sufficient threat to our national security. The greater the threat, the greater is the risk of inaction – and the more compelling the case for taking anticipatory action to defend ourselves, even if uncertainty remains as to the time and place of the enemy's attack.

Taking all of this information into consideration, one has to observe that the legality of anticipatory or pre-emptive/preventive self-defence raises considerable doubts. This is because of the vague formulation of these doctrines and the imprecise criteria that allow for the use of force in self-defence. This statement is valid regardless of which method of warfare is chosen by the state claiming the right to self-defence, whether AWs are at risk or not.

The deployment of AWs would raise considerably more doubts about the legality of anticipatory, pre-emptive or preventive self-defence. Firstly, before programming and allowing AWs to strike in anticipatory, pre-emptive or preventive self-defence, a state would have to make decisions about the parameters and boundaries of such self-defence. A number of the seminal questions to be answered are as follows:

- What is 'anticipatory', 'pre-emptive' or 'preventive'?
- When does a threat to state security require the anticipatory, pre-emptive or preventive strike?
- What kind of data should AWs analyse before deciding on the anticipatory, pre-emptive or preventive strike?
- How should AWs decide when to conduct an anticipatory, pre-emptive or preventive strike?
- If there was no armed attack, how should an AW choose the target and assess the proportionality of the attack?
- How are AWs supposed to decide that the use of force is *ultima ratio*?

Secondly, it was mentioned earlier that humans exert control over fully autonomous weapons in the development and deployment phases: when they are establishing their objectives and deciding on their activation. However, if it is impossible to pre-programme AWs with all the possible scenarios of the use of force in self-defence after an armed attack occurs, it is yet more problematic to come up with all the possible circumstances of an anticipatory, pre-emptive or preventive strike; this would make AWs yet more unpredictable than ever.

Thirdly, if there is the assumption that an AW could take the decision to strike first on the basis of a potential threat on the part of another state, the consequences of such a solution for *jus in bello* must be considered. The strike itself may amount to the launch of an armed conflict between a state that could be potentially threatened and the state that potentially may constitute a threat. However, does it matter that an AW took the decision about the strike, independently and without human intervention? To put it differently, 'can it really be said that the nation has gone to war?'<sup>92</sup> Can such action by

<sup>92</sup> Phillips (n 64) 166.

an AW be considered a manifestation of the intention of a state to start an armed conflict?<sup>93</sup> No such case has yet occurred. However, it is important to find clear answers to these questions. It seems that if the conduct of an AW may be assigned to a state, then the actions taken by the AW should be understood as state actions; in this case an AW strike should be understood as the manifestation of the intention of the state to start the conflict. Consequently, all state obligations deriving from *jus in bello* would follow after the decision taken by the AW. Only a fatal mistake or the unforeseen collapse of the system providing the data for the AW may excuse the state from such responsibility.

Thus, it is highly improbable that any state would decide to activate an AW and allow it to strike first immediately following the AW assessment that the future threat requires the use of force. A more conceivable scenario is that humans could use AWs, or systems connected with them, to evaluate whether there is a threat that requires the anticipatory use of force.<sup>94</sup> However, ultimately it would be a human who would have to authorise the anticipatory action of the AW.

## 5. Conclusions

The employment of AWs to respond to an armed attack would appear to be one of the most effective ways of using this kind of weapon. An attacked state does not have to engage human soldiers in potentially life-threatening conflicts; also the reaction of the victim state may be more rapid and precise than that calibrated only by humans.

While it is true that the deployment of AWs is both a lawful and powerful method of warfare, the constraints that derive from their employment should be noted. AWs are able to react properly and in a predictable manner only to pre-programmed scenarios; thus, the fulfilment of the criteria of necessity and proportionality can be assured only in this limited number of cases. For the same reasons, their autonomous functions should not be exploited to conduct anticipatory, pre-emptive or preventive strikes. It does not mean that the deployment of AWs must always violate international law. On the contrary, aptly pre-programmed, AWs potentially could abide by legal norms more effectively than a human. However, the legality of the use of force consists often of considerably more than just strict compliance with norms. It also requires conscious judgement, previous experience and moral compass – clearly, traits that AWs could never have.

If AWs are weapons of the future and the era of ever-more sophisticated machines is inevitably approaching, maybe it is high time to revise some traditional concepts, such as those connected with self-defence? ‘We must adjust to

<sup>93</sup> Peter Asaro, ‘On Banning Autonomous Weapon Systems: Human Rights, Automation, and the Dehumanization of Lethal Decision-Making’ (2012) 94(886) *International Review of the Red Cross* 687, 692.

<sup>94</sup> See also Nathan Leys, ‘Autonomous Weapon Systems, International Crises, and Anticipatory Self-Defense’ (2020) 45 *The Yale Journal of International Law* 377, 382.

changing times and still hold to unchanging principles'.<sup>95</sup> This phrase, once said by a former US president in the context of human rights, holds a universal truth that also perfectly addresses the relationship between ever-advancing technology and the core principles of international law. The deployment of AWs, despite being a huge military asset, may, on occasion, violate the criteria of necessity and proportionality. It is also difficult to devise a way in which AWs could legally exercise anticipatory, pre-emptive or preventive self-defence (if to assume that such a form of self-defence is legal at all). Does it mean that the international community should resign from the core international legal principles in order to allow the employment of new technology: a weapon? Even if these principles seem strict or even outdated, their aim is to keep the entire international legal community safe, and not only those states that can afford the production and use of AWs. The revision of international law, which would allow the application of new weapons, could be opening a Pandora's Box: in the future, inventors would care only about the effectiveness of weapons and not their legality, because they would be aware that the international legal framework may be adjusted in any case. Thus, as long as states are allowed to seek ways of countering new threats to their security, they need to comply with core legal principles, for the sake of international peace and security.

**Acknowledgements.** The author is grateful to Ido Rosenzweig, Director of Cyber, Belligerencies and Terrorism Research of the Minerva Center for the Study of Law under Extreme Conditions; to Yahli Shereshevsky, who served as respondent to this article during the 4th Young Researchers Workshop on Terrorism and Belligerency Human Enhancement and Advanced Technologies in Terrorism and Belligerencies, as well as to all other participants in the Conference who took part in the discussion over this article.

**Funding statement.** Not applicable.

**Competing interests.** The author declares none.

---

<sup>95</sup> President Jimmy Carter, Presidential Medal of Freedom, 1999, 'Statement on Human Rights', <https://clintonwhitehouse4.archives.gov/Initiatives/Millennium/capsule/carter.html>.

**Cite this article:** Agata Kleczkowska, 'Autonomous Weapons and the Right to Self-Defence' (2023) 56 *Israel Law Review* 24–40, <https://doi.org/10.1017/S002122372200019X>