

## *Jus in bello* and *jus ad bellum* arguments against autonomy in weapons systems: A re-appraisal

Daniele Amoroso\*

### 1. Introduction

The ethical and legal implications of the development and use of weapons systems able to perform the critical functions of target selection and engagement *autonomously* (ie without any intervention by human operators) are currently in the spotlight. The issue has recently gained widespread media coverage with the launch, on 21 August 2017, of an open letter signed so far by founders and CEOs of 126 Artificial Intelligence (AI) and robotics companies, including the ‘tech-star’ Elon Musk, who ‘implored’ States to prevent an arms race in autonomous weapons systems (AWS).<sup>1</sup>

This constitutes, however, only the latest development of a far broader debate, which was boosted in November 2012 by the publication of two coeval, and somewhat complementary documents on AWS, namely: a policy directive by the US Department of Defense on ‘Autonomy in Weapons Systems’ (US DoD Directive),<sup>2</sup> and a report by Human Rights Watch and the Harvard Law School’s International Human Rights Clinic (2012 HRW-IHRC Report)<sup>3</sup> calling for a ban on AWS,

\* Associate Professor in International Law, University of Cagliari.

<sup>1</sup> Future of Life Institute, ‘An Open Letter to the United Nations Convention on Certain Conventional Weapons’ (21 August 2017) <<https://futureoflife.org/autonomous-weapons-open-letter-2017>>. Similar concerns were voiced by a surprisingly large group of high-profile experts in the fields of robotics and Artificial Intelligence (including, among others, Stephen Hawking, Jaan Tallinn, Frank Wilczek as well as, again, Elon Musk) in a 2015 Open Letter. See Future of Life Institute, ‘Autonomous Weapons: An Open Letter From AI & Robotics Researchers’ (28 July 2015) <<https://futureoflife.org/open-letter-autonomous-weapons>>.

<sup>2</sup> US Department of Defense, ‘Autonomy in Weapons Systems’ Directive 3000.09 (21 November 2012).

<sup>3</sup> Human Rights Watch (HRW) and Harvard Law School’s International Human Rights Clinic (IHRC), ‘Losing Humanity. The Case against Killer Robots’ (19 November 2012) <[www.hrw.org/report/2012/11/19/losing-humanity/case-against-killer-robots](http://www.hrw.org/report/2012/11/19/losing-humanity/case-against-killer-robots)>.



which laid the groundwork for initiating the pro-ban campaign ‘Stop Killer Robots’.<sup>4</sup>

In the following years, thanks to staunch lobbying by NGOs joining the campaign, AWS have become an issue of common concern for States and International Organizations, which have debated the way forward with them both in formal and informal fora.<sup>5</sup> So far, the most promising venue for discussion has proven to be the institutional framework of the UN Convention on Conventional Weapons (CCW). Indeed, after a three-year cycle of informal meetings,<sup>6</sup> CCW State parties decided to establish an open-ended Group of Governmental Experts (GGE) with the mandate of exploring ‘possible recommendations on options’ for addressing *Lethal* AWS, to be submitted to the High Contracting Parties.<sup>7</sup>

The establishment of a GGE signals the readiness by CCW State parties to take the issue seriously and represents a cautious move towards the adoption of a legally binding instrument on AWS. Its first session, scheduled to be held in Geneva from 13 to 17 November 2017,<sup>8</sup> is therefore likely to represent a defining moment in the dialectics between the supporters of a ban on AWS and those who oppose it. Still, one gets the impression that, notwithstanding the sheer number of valuable publications issued by pro-ban campaigners, the latter do not fully address the criticisms raised by the opposition, leaning as they do more towards reaffirming their stances than rebutting their opponents’ counter-arguments.<sup>9</sup> Such an attitude could risk undermining the authorita-

<sup>4</sup> Stop Killer Robots, ‘Urgent Action Needed to Ban Fully Autonomous Weapons. Non-governmental organizations convene to launch Campaign to Stop Killer Robots’ (30 May 2013) <[http://stopkillerrobots.org/wp-content/uploads/2013/04/KRC\\_LaunchStatement\\_23Apr2013.pdf](http://stopkillerrobots.org/wp-content/uploads/2013/04/KRC_LaunchStatement_23Apr2013.pdf)>.

<sup>5</sup> A fairly comprehensive chronology of the initiatives taken in this respect at international and national levels is available at <[www.stopkillerrobots.org/chronology](http://www.stopkillerrobots.org/chronology)>.

<sup>6</sup> All transcripts and documents of the CCW Meetings of Experts are available at <[www.unog.ch](http://www.unog.ch)>.

<sup>7</sup> See Final Document of the Fifth Review Conference, 23 December 2016, UN Doc. CCW/CONF.V/10, Decision 1, which recalls, in turn, the Report of the 2016 Informal Meeting of Experts on Lethal Autonomous Weapons Systems (LAWS) (10 June 2016) UN Doc CCW/CONF.V/2.

<sup>8</sup> A first session was scheduled to be held from 21 to 25 August 2017. Regrettably, however, due to financial constraints, it was cancelled. This was among the triggers of the 2017 Open Letter.

<sup>9</sup> Such a flawed approach seems to emerge in the way (some) counter-arguments are



tiveness of the pro-ban campaign, which could undo the huge advocacy efforts that have led to the convening of the GGE meeting in November. And this is despite the fact that most of the concerns underlying the campaign are arguably well founded.

On these premises, the present paper will re-appraise the main pro-ban arguments, with a view to testing them against their respective counter-arguments and, where necessary, to backing them up. To this end, an attempt will be made to carry out a frank assessment of the strengths and weaknesses of the reasons put forth by both sides of the debate. As the title suggests, this analysis will be mainly focused on *jus in bello* and *jus ad bellum* issues. Although international human rights law undeniably plays a role in the pro-ban argument strategy, it will probably remain on the sidelines of the discussion in Geneva (as it did during the previous initiatives within the CCW framework). Furthermore, the human rights implications of AWS are addressed in more detail in the contribution by Andrea Spagnolo, in this Zoom-in.

The paper will be organised as follows. Section 2 will summarise the debate on what is meant by ‘autonomous’ in weapons systems and will suggest a possible way out of this definitional conundrum. In this regard, the strategic shift in focus from the notion of ‘autonomy’ to that of ‘meaningful human control’ (or lack thereof) will also be taken into account. Sections 3 to 6 will provide a re-appraisal of the main arguments in favour of a ban on AWS. In particular, section 3 will deal with the argument whereby AWS are unable to comply with some basic rules of International Humanitarian Law (IHL); section 4 will assess the contention that the deployment of AWS on the battlefield would create an accountability gap; section 5 will inquire as to whether the delegation of lethal decision-making to non-humans runs contrary to the ‘principles of humanity’ and ‘the dictates of public conscience’ under the Martens Clause; finally, section 6 will consider the claim according to which the proliferation of AWS would make wars unprecedentedly easier to wage, so jeopardising the enduring validity of the prohibition on the use of force under Article 2(4) UN Charter. In reviewing these arguments, their

addressed in a document issued by HRW and IHRC at the end of 2016: HRW and IHRC, ‘Making the Case: The Dangers of Killer Robots and the Need for a Preemptive Ban’ (9 December 2016) <[www.hrw.org/report/2016/12/09/making-case/dangers-killer-robots-and-need-preemptive-ban](http://www.hrw.org/report/2016/12/09/making-case/dangers-killer-robots-and-need-preemptive-ban)> (2016 HRW-IHRC Report).



interconnectedness and mutually reinforcing character will be highlighted: this will be spelt out in the concluding paragraph (section 7).

## 2. *Escaping the definitional conundrum: 'Autonomy' in weapons systems and meaningful human control*

Scholarly analyses on the legal implications of AWS are routinely preceded by a discussion on what makes a weapons system 'autonomous'. Methodologically, this is justified by the persistence, at the international level, of a certain disagreement as to the definition of AWS. For our purposes, there is an additional reason for addressing this issue at the outset, which lies in the fact that the definitional conundrum represents the first hurdle that the pro-ban front must overcome to make its case against AWS. Critics of a ban, indeed, alternatively argue that negotiations on the issue would be *premature* or even *pointless*, as truly autonomous weapons do not yet exist, nor will they exist in any foreseeable future;<sup>10</sup> or that they would be *infeasible* because of the lack of a shared definition of AWS.<sup>11</sup>

To start with, one should note that, while it is true that a universally shared definition of AWS is still lacking, the debate thereon is not fragmented into a myriad of competing notions, but is rather polarised around two basic construals of autonomy, well-epitomised by the definitions advanced by the UK Ministry of Defence (MoD) and the US Department of Defense (DoD), respectively.

The UK MoD definition sets a rather demanding requirement for autonomy. In a document it issued in 2011, AWS are characterised as systems 'capable of understanding higher level intent and direction', namely 'of achieving the same level of situational understanding as a human', which enables them 'to take appropriate action to bring about the desired state'.<sup>12</sup> In a similar vein, a paper by the NATO Joint Air Power Competence Centre (JAPCC) includes 'consciousness' and 'self-

<sup>10</sup> See, for instance, the Opening Statement by the UK delegation at the 2016 CCW Meeting of Experts (General Exchange, 1), whereby AWS 'do not, and may never, exist'.

<sup>11</sup> See, for references, 2016 HRW-IHRC Report (n 9) 42-43.

<sup>12</sup> UK Ministry of Defence, 'The UK Approach to Unmanned Aircraft Systems' Joint Doctrine Note 2/11 (30 March 2011) 14 (emphasis added).



determination' among the features defining truly autonomous weapons systems.<sup>13</sup> In other words, although not reaching the point of equating AWS to moral agents 'free and capable of acting on their genuine intentions',<sup>14</sup> these definitions presuppose a 'strong' AI, not far from human-like general intelligence. These requirements are hardly satisfied by any existing weapons system. Moreover, it is difficult to make reasonable predictions about the prospects of constructing machines of this kind. Therefore, the UK MoD and NATO JAPCC notions of 'autonomy' project AWS in some undetermined technological future.

The US DoD proposes a less demanding requirement on weapons systems to count as autonomous, whereby AWS must be capable, 'once activated, to select and engage targets without further intervention by a human operator'.<sup>15</sup> Quite significantly, the 2012 HRW-IHRC Report employed strikingly similar terminology, speaking of 'fully autonomous weapons that could select and engage targets without human intervention'.<sup>16</sup> A largely overlapping requirement for autonomy has also been put forth by the International Committee for the Red Cross (ICRC), which refers to 'weapons that can independently select and attack targets, ie with autonomy in the 'critical functions' of acquiring, tracking, selecting and attacking targets'.<sup>17</sup>

As one can see, these definitions allow a number of presently operating weapons systems to qualify as autonomous. These include anti-materiel defensive systems like the German *Nächstbereichschutzsystem* MANTIS<sup>18</sup> and the Israeli Iron Dome,<sup>19</sup> active protective systems for vehicles (eg the South-African/Swedish LEDS-150<sup>20</sup>), loitering weapons systems like the Israeli anti-radiation Harpy<sup>21</sup> and Harop systems,<sup>22</sup> a

<sup>13</sup> NATO JAPCC, 'Future Unmanned System Technologies. Legal and Ethical Implications of Increasing Automation' (Germany, November 2016) 9.

<sup>14</sup> G Tamburrini, 'On Banning Autonomous Weapon Systems: From Deontological to Wide Consequentialist Reasons' in N Bhuta et al (eds), *Autonomous Weapons Systems: Law, Ethics, Policy* (CUP 2016) 122, 124.

<sup>15</sup> US DoD Directive (n 2) 13-14.

<sup>16</sup> 2012 HRW-IHRC Report (n 3) 1.

<sup>17</sup> ICRC, 'Autonomous weapon systems: Technical, military, legal and humanitarian aspects. Expert meeting, Geneva, Switzerland, 26-28 March 2014' (1 November 2014) 3.

<sup>18</sup> See <[www.army-technology.com/projects/mantis/](http://www.army-technology.com/projects/mantis/)>.

<sup>19</sup> See <[www.army-technology.com/projects/irondomeairdefencemi/](http://www.army-technology.com/projects/irondomeairdefencemi/)>.

<sup>20</sup> See <<http://saab.com/land/force-protection/self-protection/leds>>.

<sup>21</sup> See <[www.iai.co.il/2013/36694-16153-en/Business\\_Areas\\_Land.aspx](http://www.iai.co.il/2013/36694-16153-en/Business_Areas_Land.aspx)>.

variety of offensive fire-and-forget munitions, like the British Brimstone missile,<sup>23</sup> and stationary (ie non-mobile) robotic sentinels, like the South-Korean Super aEgis II, which surveils the border between North and South Korea.<sup>24</sup> Accordingly, by adopting the US DoD requirement for autonomy (or alike), one could easily dismiss the argument whereby legal discussion on AWS would be premature or pointless: far from being Sci-Fi creatures, autonomous weapons are here and now.<sup>25</sup>

This does not explain, however, why a more liberal requirement should be preferred. Admittedly, indeed, US DoD-like definitions are descriptively unsatisfactory for their overinclusiveness, in that they place in the same class both static robotic sentinels located in uncluttered operational environments, like the South-Korean Super aEgis II, and any future offensive system that will be supposedly capable of operating autonomously in the fog of war. In order to introduce more accurate distinctions, future research must address the formidable problem of identifying a non-arbitrary cutting point in between the perceptual, cognitive, and action capabilities of systems like the Super aEgis II and those characterising AWS endowed with human-like skills.

Still, it is our contention that the establishment of an international regime for AWS should not be made conditional upon the solution of such a theoretical conundrum and that a broad, US DoD-like definition should be adopted. Regardless of the terminology used to label these technologies (autonomous, autonomic,<sup>26</sup> emergent intelligence weapons<sup>27</sup>), the fact remains that the development and use of weapons systems that – to paraphrase the US DoD definition – independently identify, select and engage targets raise a number of unprecedented ethical and legal issues, which require a concerted response by the international community, and whose discussion cannot be deferred to some undetermined technological future if and when the highly restrictive requirements envisaged by the UK MoD conditions will be met.

<sup>22</sup> See <[www.airforce-technology.com/projects/haroploiteringmuniti](http://www.airforce-technology.com/projects/haroploiteringmuniti)>.

<sup>23</sup> See <[www.army-technology.com/projects/brimstone](http://www.army-technology.com/projects/brimstone)>.

<sup>24</sup> See <[www.dodaam.com/eng/sub2/menu2\\_1\\_4.php#](http://www.dodaam.com/eng/sub2/menu2_1_4.php#)>.

<sup>25</sup> R Crootof, 'The Killer Robots Are Here: Legal and Policy Implications' (2015) 36 *Cardozo L Rev* 1837.

<sup>26</sup> See NATO JAPCC (n 13) 10.

<sup>27</sup> N Jain, 'Autonomous Weapons Systems: New Frameworks for Individual Responsibility' in N Bhuta et al (n 14) 303, 306.



On the other hand, the lack of a shared view as to what counts as ‘autonomous’ does not seem to create an insurmountable obstacle in the path towards the adoption of a legally binding instrument on AWS. In the current debate, in fact, NGOs, diplomats and scholars are becoming less and less interested in attaining precise operational definitions enabling one to decide, for each existing and developmental weapons system, whether it is autonomous or not. To be investigated, instead, is what kind or level of human control must on ethical and legal grounds be exercised on lethal weapons systems.

The idea whereby all weapons with lethal effects (including AWS) should be subject to a ‘meaningful human control’ (MHC) was ushered in by Article 36, a UK-based NGO.<sup>28</sup> This formula, which was refined in subsequent policy briefs, turned out to be particularly successful and soon influenced the overall campaign strategy. Significantly enough, in a 2014 report, Human Rights Watch and the Harvard Law School’s International Human Rights Clinic came to define AWS as those that ‘would identify and fire on targets without meaningful human intervention’.<sup>29</sup> The lack of meaningful human control is thereby embodied in the very definition of autonomous weapons, with the consequence that the campaign for banning AWS and that for introducing a MHC requirement end up pursuing basically overlapping goals. As has been rightly observed, ‘the notion of “meaningful human control” is just another word [...] for a partial prohibition, namely a ban on full autonomy over certain (critical) functions of a weapons system’.<sup>30</sup>

The MHC formula has rapidly attracted the interest of an increasing number of States, which have been using it as a viable starting point for the discussion on Lethal AWS at the CCW Informal Meetings of Experts.<sup>31</sup> Indeed, even the US and the UK, while not fully embracing the

<sup>28</sup> Article 36, ‘Killer Robots: UK Government Policy on Fully Autonomous Weapons’ (April 2013) <[www.article36.org/wp-content/uploads/2013/04/Policy\\_Paper1.pdf](http://www.article36.org/wp-content/uploads/2013/04/Policy_Paper1.pdf)>.

<sup>29</sup> HRW and IHRC, ‘Shaking the Foundations: The Human Rights Implications of Killer Robots’ (12 May 2014) 1 <[www.hrw.org/report/2014/05/12/shaking-foundations/human-rights-implications-killer-robots](http://www.hrw.org/report/2014/05/12/shaking-foundations/human-rights-implications-killer-robots)> (2014 HRW-IHRC Report).

<sup>30</sup> N Bhuta, S Beck, R Geiss, ‘Present Futures: Concluding Reflections and Open Questions on Autonomous Weapons Systems’ in N Bhuta et al (n 14) 347, 381.

<sup>31</sup> DA Lewis, G Blum, NK Modirzadeh, ‘War-Algorithm Accountability’ Harvard Law School Program on International Law and Armed Conflict, Research Briefing (August 2016) 62. See also United Nations Institute for Disarmament Research, *The Weaponization of Increasingly Autonomous Technologies: Considering how Meaningful Human*



notion of MHC, agree on the need to establish some level of human control over weapons systems, including autonomous ones.<sup>32</sup> This shift in focus on MHC has the undoubted advantage of setting aside the controversies surrounding the notion of ‘autonomy’ in weapons systems: No matter which weapons systems qualify as autonomous and which do not, what counts instead is that human operators should exercise a meaningful control over all of them.<sup>33</sup>

In the following sections, therefore, the main legal considerations supporting the case for a ban on lethal weapons without MHC will be expounded.<sup>34</sup>

### 3. *AWS are unable to comply with IHL principles*

A significant portion of the discussion concerning the legal implications of lethal AWS revolves around their ability to operate in compli-

*Control Might Move the Discussion Forward* (UNIDIR 2014) <[www.unidir.ch/files/publications/pdfs/considering-how-meaningful-human-control-might-move-the-discussion-forward-en-615.pdf](http://www.unidir.ch/files/publications/pdfs/considering-how-meaningful-human-control-might-move-the-discussion-forward-en-615.pdf)>.

<sup>32</sup> US DoD Directive, 2 (‘It is DoD policy that: a. Autonomous and semi-autonomous weapon systems shall be designed to allow commanders and operators to exercise *appropriate levels of human judgment* over the use of force’) (emphasis added); Article 36, ‘UK government: Defining ‘human control’ essential at killer robots talks in 2017’ (12 January 2017) <[www.article36.org/autonomous-weapons/uk-govt-response-2017](http://www.article36.org/autonomous-weapons/uk-govt-response-2017)>.

<sup>33</sup> Of course, it is open to question under what conditions human control is to be deemed as meaningful. Unlike the discussion on the notion of ‘autonomy’, however, this does not raise a (possibly unsolvable) theoretical problem; rather, it constitutes a normative issue to be addressed precisely through negotiations. One should note, in this respect, that pro-ban campaigners are rather flexible as to the terminology to be employed. See Article 36, ‘Key elements of meaningful human control’ (April 2016) <[www.article36.org/wp-content/uploads/2016/04/MHC-2016-FINAL.pdf](http://www.article36.org/wp-content/uploads/2016/04/MHC-2016-FINAL.pdf)> (‘There are other words that could be used instead of “meaningful”, for example: appropriate, effective, sufficient, necessary’).

<sup>34</sup> It has been suggested elsewhere, in a paper co-authored with Guglielmo Tamburrini, that the same (ethical and) legal considerations supporting the need for a MHC requirement on weapons systems may provide sound guidance as to the level of control that humans have to exercise over them. See D Amoroso, G Tamburrini, ‘The Ethical and Legal Case against Autonomy in Weapons Systems’ (2017) *Global Jurist* (22 September 2017) <[www.degruyter.com/view/j/gj.ahead-of-print/gj-2017-0012/gj-2017-0012.xml?format=INT](http://www.degruyter.com/view/j/gj.ahead-of-print/gj-2017-0012/gj-2017-0012.xml?format=INT)>.





ance with IHL principles of distinction, military necessity and proportionality, *at least as well as a competent and conscientious human soldier*.<sup>35</sup>

Supporters of a ban contend, in the first place, that AWS are unlikely to be endowed with the capability of distinguishing between civilians and combatants in erratic and surprise seeking warfare environments, and even less with that of establishing whether civilians have lost protection from attacks because they ‘directly participated in hostilities’ or assumed ‘continuous combat functions’.<sup>36</sup> Since the distinction between combatants and protected persons is no longer based on easily perceivable distinctive signs such as military uniforms, but rather on people’s ‘behaviour and actions on the battlefield’,<sup>37</sup> an IHL-compliant AWS should possess a variety of perceptual and cognitive abilities that are hardly attainable in any foreseeable future.<sup>38</sup> This includes viewpoint-independent recognition of bodily postures and gestures in the variable perceptual conditions of unstructured warfare scenarios, an understanding of emotional expressions, and real-time reasoning about deceptive intentions and actions.

Secondly, and for very similar reasons, it is doubted that AWS will ever be able to comply reliably with the prohibition, stemming from the principle of military necessity, of attacking persons rendered *hors de*

<sup>35</sup> C Heyns, ‘Report by the Special Rapporteur on extrajudicial, summary or arbitrary executions’ UN Doc A/HRC/23/47 (9 April 2013) paras 63-74. See also, among others, N Sharkey, ‘The Evitability of Autonomous Robot Warfare’ (2012) 94 *Intl Rev of the Red Cross* 787; D Akerson, ‘The Illegality of Offensive Lethal Autonomy’ in D Saxon (ed) *International Humanitarian Law and the Changing Technology of War* (Brill 2013) 65; C Grut, ‘The Challenge of Autonomous Lethal Robotics to International Humanitarian Law’ (2013) 18 *J Conflict & Security L* 5; J van den Boogaard, ‘Proportionality and Autonomous Weapons Systems’ (2015) 6 *J Intl Humanitarian L Studies* 247; K Egeland, ‘Lethal Autonomous Weapon Systems under International Humanitarian Law’ (2016) 85 *Nordic J Intl L* 89.

<sup>36</sup> On these two concepts, which can hardly be framed into clear-cut definitions, see N Melzer, *Interpretive Guidance on the Notion of Direct Participation in Hostilities under International Humanitarian Law* (ICRC 2009).

<sup>37</sup> NATO JAPCC (n 13) 21.

<sup>38</sup> P Alston, ‘Lethal Robotic Technologies: The Implications for Human Rights and International Humanitarian Law’ (2011) 21 *J of Law, Information and Science* 35, 54; Sharkey (n 35) 788-789; Grut (n 35) 11-12; R Geiss, ‘The International-Law Dimension of Autonomous Weapons Systems’, Friedrich Ebert Stiftung Study (October 2015) 14 <<http://library.fes.de/pdf-files/id/ipa/11673.pdf>>; WH Boothby, ‘Autonomous Attack – Opportunity or Spectre?’ (2013) 16 *YB Intl Humanitarian L* 71, 79-80.

*combat*.<sup>39</sup> Indeed, the recognition of behaviours that convey hostile or surrender messages and fighting incapacitation information will pose no less insurmountable challenges for AWS programmers and developers.<sup>40</sup>

Things become even trickier, and here we get to the third point, when ‘the terrible and impossible problem of proportionality’<sup>41</sup> comes into play. As is well known, the proportionality principle requires striking a balance between military gains expectedly deriving from some given course of action and harm to civilians ensuing from it.<sup>42</sup> The content of this principle, therefore, cannot be defined *in abstracto*, but needs to be tailored to the specific circumstances of the case, on the basis of a delicate balancing activity entrusted to those involved in the planning and execution of the attack.<sup>43</sup> Proportionality analysis, the argument goes, heavily relies on qualitative elements and open-textured standards such as that of the ‘reasonable military commander’, which are simply unsuitable for being coded into an algorithm.<sup>44</sup>

On the opposite side of the debate, two main arguments have been put forth in order to counter these concerns.

On the one hand, the pro-ban campaign would wrongly assume that issues of distinction and proportionality arise in every battlespace. On the contrary, it is well possible to conceive of a variety of scenarios where civilians or civilian objects are totally absent (eg a battleship in the high seas). In these contexts, IHL would not pose any obstacle to

<sup>39</sup> On which see J-M Henckaerts, L Doswald-Beck, *Customary International Humanitarian Law* (vol 1, CUP 2006) 164-170.

<sup>40</sup> R Sparrow, ‘Twenty Seconds to Comply: Autonomous Weapon Systems and the Recognition of Surrender’ (2015) 91 *Intl L Studies* 699.

<sup>41</sup> GD Solis, *The Law of Armed Conflict: International Humanitarian Law in War* (CUP 2016) 293.

<sup>42</sup> See, generally, E Cannizzaro, ‘Proportionality in the Law of Armed Conflict’ in A Clapham, P Gaeta (eds), *The Oxford Handbook of International Law in Armed Conflict* (OUP 2014) 332.

<sup>43</sup> For this reason, proportionality has been described as a ‘secondary process of law-making’ (ibid 332).

<sup>44</sup> Sharkey (n 35) 789-790; Heyns (n 35) paras 70-73; Grut (n 35) 12-14; Egeland (n 35) 103-105; M Wagner, ‘The Dehumanization of International Humanitarian Law: Legal, Ethical, and Political Implications of Autonomous Weapon Systems’ (2014) 47 *Vanderbilt J Transnational L* 1371, 1398-1399; Boothby (n 38) 82-83; P Kalmanovitz, ‘Judgment, Liability and the Risks of Riskless Warfare’ in N Bhuta et al (n 14) 145, 151-152.



the deployment of AWS, provided that all other legal requirements are respected (for instance, the system is not loaded with chemical or bacteriological weapons or with weapons otherwise aimed at causing ‘superfluous injury or unnecessary suffering’).<sup>45</sup>

This is certainly correct, insofar as the principles of distinction and proportionality are concerned. It does not seem to consider, however, that the rule prohibiting attacks against enemies *hors de combat*, which also raises perceptive and evaluative challenges for AWS, is virtually applicable in every warfare scenario, regardless of whether civilians are present or not. Lethal uses of AWS, in other words, remain problematic, no matter what the circumstances of their deployment are.

It could be further counter-argued, at this juncture, that the pro-ban front sets an exceedingly high threshold for AWS, which goes well beyond what is normally required in relation to human-controlled attacks.<sup>46</sup> In fact, neither high-altitude aerial bombings nor ‘beyond visual range’ attacks have been considered unlawful *per se*,<sup>47</sup> notwithstanding the fact that the operator has limited methods of ascertaining whether human targets are defenceless or surrendering. This observation has its fair share of merits as well. Yet, it fails to take into account that if a military operator launches an attack against enemies *hors de combat*, he or she may be held criminally accountable for this behaviour – something that is far more difficult in case of attacks launched by AWS. This raises an ‘accountability gap’ issue, which I will come back later.<sup>48</sup>

On the other hand, the Campaign would base its IHL arguments on unverifiable factual claims as to what AWS will be able (or will never be able) to do. As challenging as it may appear, it cannot be ruled out that one day AWS will pass an ‘IHL Turing test’, under which an objective observer would not be able to distinguish a machine from a human on the sole basis of their compliance with IHL principles.<sup>49</sup> Again, this is a

<sup>45</sup> MN Schmitt, ‘Autonomous Weapon Systems and International Humanitarian Law: A Reply to the Critics’ (2013) 1 Harvard National Security J Features 1, 11. See also van Boogard (n 35) 262.

<sup>46</sup> Schmitt (n 45) 12.

<sup>47</sup> With particular regard to high-altitude aerial bombings, see Final Report to the Prosecutor by the Committee Established to Review the NATO Bombing Campaign Against the Federal Republic of Yugoslavia (13 June 2000) para 56.

<sup>48</sup> See Section 4.

<sup>49</sup> K Anderson, MC Waxman, ‘Debating Autonomous Weapon Systems, Their Ethics, and Their Regulation Under International Law’ in R Brownsword, E Scotford,

convincingly argued point. For all we know, lack of IHL compliance is a contingent situation that may no longer hold at some future time. Still, some caveats are in order.

Firstly, while the possibility that IHL principles are effectively translated into machine algorithms cannot be radically denied, one should carefully note that ‘translation’ does not mean ‘oversimplification’. As an example of a scholarly trend towards oversimplification, one may mention an influential article by Michael N Schmitt, where it is suggested that human commanders could ‘pre-programme’ AWS to carry out proportionality analysis by assigning a certain value in terms of collateral damage to each military objective (eg a tank or a military base).<sup>50</sup> Indeed, this completely overlooks the fact that the proportionality principle must be abided by not only in the planning (viz pre-programming) and decision phases of an attack, but also throughout its execution.<sup>51</sup> Notably, proportionality has to be re-assessed in the light of the circumstances of the case, which may well have changed since the attack was launched. Therefore, the ‘value’, in terms of acceptable collateral damage assigned to military objectives cannot be determined once and for all, but has to be constantly adjusted to match the shifting context where the AWS has been deployed. Should those adjustments be carried out in real time by human operators, weapons systems would hardly qualify as autonomous: after all, they would be subject to an (arguably) meaningful human control. If, on the other hand, AWS were to be endowed with the capability to effect proportionality assessments throughout the execution phase, their algorithms clearly would need to be far more sophisticated than the one envisaged by Schmitt.

Secondly, it should be made sufficiently clear that the development of an AWS able to pass the ‘IHL Turing test’ in cluttered scenarios, while possible *in abstracto*, poses a number of formidable technological challenges, whose solution ‘does not represent the expected outcome of an existing research programme’,<sup>52</sup> but rather constitutes a long-term

K Yeung (eds), *The Oxford Handbook of Law, Regulation, and Technology* (OUP 2017) 1097, 1109-1110.

<sup>50</sup> Schmitt (n 45) 20-21.

<sup>51</sup> E Liebllich, E Benvenisti, ‘The Obligation to Exercise Discretion in Warfare: why Autonomous Weapons Systems are Unlawful’ in N Bhuta et al (n 14) 245, 269-275.

<sup>52</sup> Bhuta, Beck, Geiss (n 30) 351.



objective speculated by (often self-professed visionary) roboticists.<sup>53</sup> In the absence of ‘black swan’ events, therefore, the actual making of IHL-compliant AWS is highly uncertain and, in any case, to be temporally located in a very distant future.

Thirdly and finally, even if one accepts that one day AWS might be capable of human-like performances with respect to adherence to IHL, this says nothing as to whom will be held accountable for AWS’ breaches of IHL (in fact, human like – and even ‘higher-than-human’ – performances are not equal to a spotless application of IHL standards), or as to whether non-human lethal decision-making, regardless of its compliance with IHL, conforms to the ‘principles of humanity’ and the ‘dictates of public conscience’ under the Martens Clause. These constitute additional reasons in favour of a ban on AWS, which will now be addressed.

#### 4. *The deployment of AWS will determine accountability gaps*

An argument often raised against autonomy in weapons systems is that it is bound to create accountability gaps. Indeed, even the more convinced proponents of AWS are compelled to admit that, no matter how accurate, these systems will never be completely flawless.<sup>54</sup> As a consequence, it is quite possible for an AWS to act in breach of IHL and, in the most serious hypotheses, to commit acts amounting – at least materially – to war crimes.<sup>55</sup> But then who will be personally accountable for these conducts? Since AWS obviously cannot be held responsible as direct perpetrators, responsibility for their actions must be traced back to some persons in the decision-making chain.

At the outset, one should take note that the list of potentially responsible individuals is quite long, as it includes ‘the software programmer, [the manufacturer of the AWS,<sup>56</sup> the procurement official<sup>57</sup>],

<sup>53</sup> Tamburrini (n 14) 134.

<sup>54</sup> P Scharre, ‘Autonomous Weapons and Operational Risk’, Ethical Autonomy Project (February 2016) 17 <[www.cnas.org/publications/reports/autonomous-weapons-and-operational-risk](http://www.cnas.org/publications/reports/autonomous-weapons-and-operational-risk)>.

<sup>55</sup> R Crootof, ‘War Torts: Accountability for Autonomous Weapons’ (2016) 164 U Pennsylvania L Rev 1347, 1375-1377.

<sup>56</sup> Jain (n 27) 321-324.



the military commander in charge of the operation, the military personnel that sent the AWS into action or those overseeing its operation, the individual(s) who conducted the weapons review, or political leaders'.<sup>58</sup> Such a long list, far from facilitating the task of identifying the responsible individuals, is likely to raise the familiar 'many hands' problem. This problem commonly occurs in software-related accidents,<sup>59</sup> where a group of people can be held collectively responsible for a determined outcome, whereas often none of them can be individually blamed for it.<sup>60</sup>

An additional source of unaccountability lies in the fact that AWS are designed to behave in a way that cannot be predicted by its users. To be clear, unpredictability would not feature solely futuristic weapons endowed with strong AI, but also less sophisticated systems meeting the more liberal requirements for autonomy set forth by the US DoD. Although acting on the basis of goals and rules determined by humans, they have a certain autonomy as to how to accomplish a given task and, at least in some cases, machine-learning capabilities, which may allow them to learn from past errors and act differently. All of this, coupled with a dynamic and unstructured environment, may render it virtually impossible to foresee every individual action of AWS.<sup>61</sup>

How this impacts on individual accountability is easy to grasp. Clearly, there may be uncontroversial cases such as that of a machine which is deliberately pre-programmed to carry out international crimes or that of a commander who deploys an AWS in a context different from the one it was designed for, and where it was likely to commit war

<sup>57</sup> GS Corn, 'Autonomous Weapons Systems: Managing the Inevitability of "Taking the Man Out of the Loop"' in N Bhuta et al (n 14) 209, 230-238.

<sup>58</sup> M Wagner, 'Autonomous Weapon System' in R Wolfrum (ed), *Max Planck Encyclopedia of Public Intl L* (OUP 2016) margin n 21.

<sup>59</sup> H Nissenbaum, 'Accountability in a Computerized Society' (1996) 2 *Science and Engineering Ethics* 25.

<sup>60</sup> On the many hands problem, see generally DF Thompson, 'Moral Responsibility of Public Officials: The Problem of Many Hands' (1980) 74 *The American Political Science Rev* 905. See also, with specific regard to AWS, M Noorman, DG. Johnson, 'Negotiating Autonomy and Responsibility in Military Robots' (2014) 16 *Ethics and Information Technology* 51, 61; H-Y Liu, 'Refining Responsibility: Differentiating Two Types of Responsibility Issues Raised by Autonomous Weapons Systems' in N Bhuta et al (n 14) 325, 331.

<sup>61</sup> A Krishnan, *Killer Robots: Legality and Ethicality of Autonomous Weapons* (Routledge 2009) 58; Tamburrini (n 14) 127-129; Jain (n 27) 306-307; Liu (n 60) 330.



crimes.<sup>62</sup> In the majority of conceivable scenarios, however, the complexities of AWS technologies and their behavioural unpredictability in partially structured or unstructured warfare scenarios are likely to afford a powerful defence against criminal prosecution. Indeed, in most cases it would be impossible to ascertain the existence of the mental element (intent, knowledge or recklessness), which is required under international criminal law (ICL) to ascribe criminal responsibility.<sup>63</sup> As a consequence, it would be highly probable that no one person would be held criminally liable, notwithstanding that the conduct in question were to undeniably amount to a war crime.<sup>64</sup>

While generally acknowledging that AWS pose unprecedented challenges in relation to individual criminal accountability for war crimes, those sceptical of a ban have marshalled a number of alternative solutions short of a general prohibition on lethal uses of AWS.

Some have argued that no accountability gap would arise in relation to the use of AWS, by relying on the doctrine of ‘command responsibility’. It is submitted, in particular, that the commander (or the civilian supervisor) will be held criminally responsible for war crimes perpetrated by an AWS if he (or she) knew, or should have known, that it was committing or about to commit such crimes and, that notwithstanding, failed to prevent their commission or punish those responsible for it.<sup>65</sup> However, the doctrine of ‘command responsibility’, as it currently stands, is of little help. One of its constitutive elements is the commander’s knowledge of the subordinate’s behaviour, or at least its predictability. Yet, this element can hardly be applied here. On the one hand, as we have already seen, there is reason to maintain that AWS may well take unforeseeable courses of action. On the other hand, AWS’ faster-than-human reaction times would make a commander’s control to a large extent a purely speculative option.<sup>66</sup>

<sup>62</sup> Crootof (n 55) 1377.

<sup>63</sup> Wagner (n 44) 1405-1406.

<sup>64</sup> R Sparrow, ‘Killer Robots’ (2007) 24 *Journal of Applied Philosophy* 62; Grut (n 35) 14-17; HRW and IHRC, ‘Mind the Gap: The Lack of Accountability for Killer Robots’ (9 April 2015) 19-20 <[www.hrw.org/report/2015/04/09/mind-gap/lack-accountability-killer-robots](http://www.hrw.org/report/2015/04/09/mind-gap/lack-accountability-killer-robots)>; Geiss (n 38) 20-21; Egeland (n 35) 110-112; Crootof (n 55) 1375-1378.

<sup>65</sup> Schmitt (n 45) 33; NATO JACPP (n 13) 29.

<sup>66</sup> Crootof (n 55) 1380.



In order to address this problem, it has been suggested that there be a shift in the accountability focus from the deployment to the development/procurement phase. Only at this stage, would it be possible to ensure that AWS are effectively equipped with all the cognitive and evaluative capabilities needed to faithfully respect IHL principles. Accordingly, responsibility for AWS' war crimes should primarily lie with 'military procurement managers, weapons developers and legal advisors'.<sup>67</sup> In this regard, it could be argued that, by limiting the accountability focus to the development/procurement phase, one does not overcome the predictability issue set out above. If deployed in a dynamic environment, an AWS endowed is capable of taking courses of action whose reason may be unfathomable 'even to the system's designers'.<sup>68</sup> Under these circumstances, it seems highly unlikely that the officers involved in the procurement/development phase possess the culpable state of mind required to establish criminal responsibility.

The aforementioned difficulties in establishing the existence of the required *mens rea*, even in the loose form of command responsibility, have led some authors to suggest that its threshold should be lowered. It is worth recalling, in this regard, the proposal to introduce 'opaque recklessness' as a culpable state of mind in relation to AWS-related war crimes.<sup>69</sup> Under this category, the defendant would be considered to have acted with recklessness where he or she 'knows his or her conduct is risky but either fails to realize or consciously disregards the specific reasons for the riskiness'.<sup>70</sup> This would in fact allow holding the 'commander/field officer/deploying soldier' criminally accountable for AWS' war crimes, even if he or she was 'unaware of the exact risk of harm posed by the AWS's conduct' and even if the latter's actions were 'uncertain and unpredictable', provided that he or she was aware that there was 'a substantial and unjustified risk' of some unspecified 'dangerous occurrence'.<sup>71</sup>

This proposal, like others of the same kind, comes across as particularly problematic. In dealing with accountability issues, one should always take care not to confuse the fight against impunity with 'scape-

<sup>67</sup> Corn (n 57) 230-238.

<sup>68</sup> Crootof (n 55) 1373. See also Wagner (n 44) 1402-1403.

<sup>69</sup> Jain (n 27) 317.

<sup>70</sup> *ibid.*

<sup>71</sup> *ibid* 318.



goat[ing] proximate human beings'.<sup>72</sup> And indeed, to the extent that it stretches the notion of culpability up to the outer limits of strict liability, the notion of 'opaque recklessness' seems to be precisely geared towards that alarming direction. Furthermore, it should be kept in mind that any lowering of the *mens rea* threshold would entail a reform of international criminal law as it currently stands. While this is not an evil *per se*, one should be aware that States would be no less reluctant to relax the requirements for holding their soldiers responsible for war crimes than to negotiate a ban on lethal uses of AWS.

One could object to the above, by arguing that unpredictability, just like inability to comply with IHL principles, is not an inherent feature of AWS. Even in this case, it cannot be ruled out that technological advances will make it possible to reliably predict the courses of action undertaken by an AWS, so providing a solid basis for ascribing responsibility to the individuals involved in its development and deployment.<sup>73</sup> Admittedly, that could be very challenging and could perhaps only be achieved in a distant future, but it nonetheless falls within the realms of possibility. Still, such an objection does not detract from the conclusion that AWS are bound to create accountability gaps. This is because, with the deployment of these systems on the battlefield, a category of culpable agents would be (almost) completely swept away: direct perpetrators. Except for in the highly unlikely case of AWS being programmed for the purpose of committing war crimes or being commanded to do so, accountability for AWS-related war crimes will necessarily take the form of vicarious responsibility.<sup>74</sup> And it is legitimate to doubt whether this could fully capture the seriousness of the crimes that could be committed by AWS.

The need to provide a proper legal response to serious violations of international law such as war crimes also explains, *a fortiori*, why proposals to fill the accountability gap by relying on forms of collective responsibility (such as State responsibility<sup>75</sup> or corporate product liabil-

<sup>72</sup> Liu (n 60) 341.

<sup>73</sup> *ibid* 335

<sup>74</sup> *ibid* 339-341.

<sup>75</sup> See, for instance, Crootof (n 55); Anderson, Waxman (n 49) 1110; NATO JACPP (n 13) 39; DN Hammond, 'Autonomous Weapons and the Problem of State Accountability' (2015) 15 *Chicago J Intl L* 652.

ity<sup>76</sup>) are largely unsatisfactory. In fact, the crucial, two-fold function of *detering* the commission of international crimes and adequately *retributing* the offender for the harm done is peculiar to ICL and cannot be performed in the same way by collective responsibility, for the well-known reason that international crimes ‘are committed by men, not by abstract entities, and only by punishing individuals who commit such crimes can the provisions of international law be enforced’.<sup>77</sup>

##### 5. AWS’ lethal decision-making runs contrary to the Martens Clause

Since the pro-ban campaign’s inception, the Martens Clause has been included in the array of arguments in favour of a ban on lethal AWS. The history and content of the Clause are well known to international lawyers. It made its first appearance in international legal parlance in 1899, when it was inserted – on the proposal of the Russian publicist Fyodor Fyodorovich Martens (from whom it takes its name) – in the Preamble of the Second Hague Convention containing the Regulations on the Laws and Customs of War on Land. Subsequently incorporated in a number of IHL instruments (including the CCW), in its modern formulation the Martens Clause states that, absent any specific regulation, ‘the civilian population and the combatants shall at all times remain under the protection and authority of the principles of international law derived from established custom, from the principles of humanity and from the dictates of public conscience’.<sup>78</sup>

<sup>76</sup> NATO JACPP (n 13) 29-30. See also, although more problematically, Jain (n 27) 323-324.

<sup>77</sup> International Military Tribunal Nuremberg, *The Trial of German Major War Criminals* (Judgment, 1 October 1946). Proceedings of the International Military Tribunal sitting at Nuremberg, Germany, Part 22 (22nd August 1946 to 1st October 1946) 447. The Norwegian delegate at the 2016 CCW Meeting of Experts put the problem in very similar terms: ‘Another intrinsic challenge with fully autonomous weapons would be ensuring individual [...] responsibility for unlawful acts in times of armed conflict. This is a cornerstone of modern international law. Without accountability, deterring and preventing international crimes becomes all that much harder’ (Statement at the Panel on ‘Challenges to International Humanitarian Law’).

<sup>78</sup> See, generally, the seminal work by Antonio Cassese, ‘The Martens Clause: Half a Loaf or Simply Pie in the Sky?’ (2000) 11 *Eur J Intl L* 187.



In the context of this campaign in particular, the Martens Clause has been invoked to contend that the deployment of weapons systems enabled to take life-or-death decisions without human supervision would run contrary to international law as stemming from both ‘the principles of humanity’ and ‘the dictates of public conscience’.<sup>79</sup>

Reliance on the ‘principles of humanity’ overlaps with the view, grounded in international human rights law, whereby the delegation of lethal decision-making to a machine would be prohibited under the principle of human dignity.<sup>80</sup> This line of argumentation is more far-reaching than the previous ones, as it is built upon principles, those of humanity and human dignity, which are both foundational and open-textured. While the basic claim underlying this argument is straightforward (‘there is a violation of the principles of humanity/human dignity when a machine kills a human being’), it is informative to distinguish analytically between two variants.

The first variant moves from the assumption that the action of suppressing a human life is legally justifiable only if it is non-arbitrary, namely by being based ‘on a considered and informed decision’.<sup>81</sup> In order to be non-arbitrary, and here it is where the principles of humanity come in, the act of killing must be grounded on *human* judgement, for only human decision-making guarantees the full appreciation ‘of the value of individual life [and] the significance of its loss’.<sup>82</sup> The second variant was formulated by the Special Rapporteur on extrajudicial, summary or arbitrary executions, Christof Heyns.<sup>83</sup> In his view, human

<sup>79</sup> 2012 HRW-IHRC Report (n 3) 35-36.

<sup>80</sup> The intersection between the two issues was underlined by Patrick Lin, in a presentation delivered at the 2015 CCW Meeting of Experts (‘The right to life and the Martens Clause’ <[www.unog.ch/80256EE600585943/\(httpPages\)/6CE049BE22EC75A2C1257C8D00513E26?OpenDocument](http://www.unog.ch/80256EE600585943/(httpPages)/6CE049BE22EC75A2C1257C8D00513E26?OpenDocument)>). In that context, the argument whereby AWS’ lethal decision-making would breach the principle of human dignity was put forth with particular vigour by Germany (2014 CCW Meeting of Experts, Opening Statement) and the Holy See (2015 CCW Meeting of Experts, Statement at the Panel on ‘Overarching Issues’).

<sup>81</sup> P Asaro, ‘On Banning Autonomous Weapon Systems: Human Rights, Automation, and the Dehumanization of Lethal Decision-making’ (2012) 94 Intl Rev of the Red Cross 687, 689.

<sup>82</sup> 2014 HRW-IHRC Report (n 29) 3.

<sup>83</sup> C Heyns, ‘A Human Rights Perspective on Autonomous Weapons in Armed Conflict: The Rights to Life and Dignity’ in R Geiss (ed) *Lethal Autonomous Weapons*

dignity would be blatantly denied if people were subject to lethal robotic decision-making, because this would place them in a position where they 'have no venue, futile or not of appealing to the humanity of the enemy'. Indeed, the decision to kill or not would be taken on the basis of hypotheticals set out in advance in the AWS programming phase, or developed by the machine itself as rules of behaviour extrapolated from its past experience. The ensuing death-or-life decision could hardly be overridden when the AWS is about to actually release force, with the consequence that the human target would be somehow 'written off' without the (even the slightest) hope of changing his/her fate.<sup>84</sup>

The main criticism levelled against this argument, in both its variants, is that it distils a number of implications from the principles of humanity and human dignity that remain ultimately undemonstrated. On the one hand, the assumption whereby, to be non-arbitrary, the use of lethal force during armed conflicts should result from value judgements that are based on human reason has been contested. Legally speaking, in fact, what matters is only that targeting decisions, regardless of whom ultimately takes them, comply with the objective requirements of IHL.<sup>85</sup> On the other hand, it has been pragmatically argued that, from the perspective of the recipients of lethal force, it is totally immaterial whether 'the threat they are exposed to comes from manned or unmanned weapons', being highly questionable that 'the mere potentiality of a human commander's mercy or compassion [would] make a difference if, in fact, this potentiality does not materialize'.<sup>86</sup>

These criticisms partially hit the mark, as they unveil how the arguments based solely on the 'principles of humanity' prong of the Clause are built on a set of *a priori* assumptions that fail the test of closer scrutiny. What these counter-arguments do not seem to take into account,

Systems Technology, Definition, Ethics, Law & Security (German Federal Foreign Office 2017) 148, 156.

<sup>84</sup> *ibid.* In a similar vein, M Brehm, 'Defending the Boundary: Constraints and Requirements on the Use of Autonomous Weapon Systems Under International Humanitarian and Human Rights Law' Geneva Academy of International Law and Human Rights, Briefing no 9 (1 May 2017) 65.

<sup>85</sup> M Sassòli, 'Autonomous Weapons and International Humanitarian Law: Advantages, Open Technical Questions and Legal Issues to be Clarified' (2014) *Intl L Studies* 308, 318.

<sup>86</sup> D Birnbacher, 'Are Autonomous Weapon Systems a Threat to Human Dignity?' in N Bhuta et al (n 14) 105, 121.



however, is that humanity and human dignity are not immutable notions, from which one could infer what counts as inhumane (or undignified) exclusively by way of logical deduction, but they are (also) ‘a function of contemporary social understandings’.<sup>87</sup>

This is well captured by the reference, in the Martens Clause, to the ‘dictates of public conscience’, which brings the ethereal ‘principles of humanity’ back to an earthly dimension, rooted in the reactions of the social basis of the international community to certain means and methods of warfare.<sup>88</sup> Admittedly, this prong of the Clause is the most controversial, as it is far from certain *whose* conscience should be taken into consideration.<sup>89</sup> Scholarly and State opinions in this respect are quite varied, as they range from those simply equating ‘public conscience’ with *opinio iuris* to those arguing that the Martens Clause elevates public opinion to a source of IHL.<sup>90</sup> Nor did the ICJ provide clearer guidance when it dealt with the Martens Clause in the *Nuclear Weapons Advisory Opinion*.<sup>91</sup>

Of course, this is not the place to engage in a thorough discussion of the issue. Nevertheless, one cannot help but notice that the idea whereby machines should not take life-or-death decisions has been gaining growing consensus within the international community at large. Evidence of this may be found, in particular, in the declarations rendered by States at the Human Rights Council in reaction to the presentation of the Heyns Report on Lethal AWS,<sup>92</sup> at the UN General Assembly First Committee on Disarmament and International Security,<sup>93</sup> and during CCW Informal Meetings of Experts;<sup>94</sup> in parliamentary initiatives

<sup>87</sup> R Sparrow, ‘Robots and Respect: Assessing the Case Against Autonomous Weapon Systems’ (2016) 30 *Ethics & Intl Affairs* 93, 109.

<sup>88</sup> R Ticehurst, ‘The Martens Clause and the Laws of Armed Conflict’ (1997) 37 *Intl Rev of the Red Cross* 125, 134.

<sup>89</sup> T Meron, ‘The Martens Clause, Principles of Humanity, and Dictates of Public Conscience’ (2000) 94 *AJIL* 78, 85.

<sup>90</sup> *ibid* 83-85.

<sup>91</sup> *Legality of the Threat or Use of Nuclear Weapons* (Advisory Opinion) [1996] ICJ Rep 226. See Cassese (n 78) 211.

<sup>92</sup> Heyns (n 35). See, for references, Stop Killer Robots, ‘Consensus: killer robots must be addressed’ (30 May 2013) <[www.stopkillerrobots.org/2013/05/nations-to-debate-killer-robots-at-un/](http://www.stopkillerrobots.org/2013/05/nations-to-debate-killer-robots-at-un/)>.

<sup>93</sup> See the chronology of the campaign at <[www.stopkillerrobots.org/chronology/](http://www.stopkillerrobots.org/chronology/)>.

<sup>94</sup> See the Appendixes to the Report by Lewis, Blum, Modirzadeh (n 31) 136-137, 146-147, and 151-223.

specifically addressing this matter;<sup>95</sup> in reports issued by international human rights supervisory bodies;<sup>96</sup> in the (qualified) criticism voiced in the Open Letters signed in 2015 and 2017, respectively, by renowned experts in the fields of robotics and Artificial Intelligence and founders and CEOs of artificial intelligence (AI) and robotics companies;<sup>97</sup> as well as in opinion surveys showing a spread of hostility to non-human lethal decision-making<sup>98</sup> (especially – and this seems worthy of note – among members of the armed forces).<sup>99</sup> While it would be highly speculative to draw any final conclusion, therefore, the emergence of a global trend in favour of banning AWS can hardly be denied.

But, even if such a trend consolidated into ‘dictates of public conscience’, what normative consequences would practically ensue? According to one author, this would be completely irrelevant, since the Martens Clause comes into play only in the absence of more specific regulation, while weapons law is currently replete with norms and principles whose application may be easily extended to AWS.<sup>100</sup> This line of argument, which ends up denying any relevance to the Martens Clause in contemporary international law, curiously resembles the one advanced by the Russian Federation before the ICJ during the advisory

<sup>95</sup> These include a resolution of the European Parliament calling on ‘the High Representative for Foreign Affairs and Security Policy, the Member States and the Council to: [...] (d) ban the development, production and use of fully autonomous weapons which enable strikes to be carried out without human intervention’. See European Parliament, ‘Resolution on the Use of Armed Drones’ (27 February 2014) 2014/2567(RSP) para 2. For other parliamentary initiatives see <[www.stopkillerrobots.org/2017/04/parliaments/](http://www.stopkillerrobots.org/2017/04/parliaments/)>.

<sup>96</sup> Heyns (n 35). See also M Kiai, C Heyns, ‘Joint report of the Special Rapporteur on the rights to freedom of peaceful assembly and of association and the Special Rapporteur on extrajudicial, summary or arbitrary executions on the proper management of assemblies’, UN Doc A/HRC/31/66 (4 February 2016); and African Commission on Human and Peoples’ Rights, ‘General Comment No 3 on the African Charter on Human and Peoples’ Rights: The Right to Life (Article 4)’, 57th Ordinary Session 4-18 November 2015 para 35.

<sup>97</sup> See above n 1 and accompanying text.

<sup>98</sup> See, most recently, Ipsos, ‘Three in Ten Americans Support Using Autonomous Weapons’ (7 February 2017) <[www.ipsos.com/en-us/news-polls/three-ten-americans-support-using-autonomous-weapons](http://www.ipsos.com/en-us/news-polls/three-ten-americans-support-using-autonomous-weapons)>.

<sup>99</sup> This is evidenced, in particular, in a survey conducted by YouGov America on the US public. See C Charpenter, ‘US Public Opinion on Autonomous Weapons’ (June 2013) <[http://duckofminerva.dreamhosters.com/wp-content/uploads/2013/06/UMass-Survey\\_Public-Opinion-on-Autonomous-Weapons.pdf](http://duckofminerva.dreamhosters.com/wp-content/uploads/2013/06/UMass-Survey_Public-Opinion-on-Autonomous-Weapons.pdf)>.

<sup>100</sup> Schmitt (n 45) 32.





proceedings in the *Nuclear Weapons* case. On that occasion, indeed, it was contended that, following the adoption of a complete code of war with the Geneva Conventions and the Protocols thereto, the Clause should 'formally be considered inapplicable'.<sup>101</sup> Yet, the ICJ openly dismissed this view, by stressing the on-going relevance of the Clause as an 'effective means of addressing the rapid evolution of military technology'.<sup>102</sup> This *dictum* applies all the more to emerging technologies such as AWS, whose disruptive features prevent a mechanical application of principles developed in relation to human-controlled weapons.

Still, widespread aversion, even if crystallised into the 'dictates of public conscience', cannot *ipso facto* outlaw the development and use of AWS. As Theodor Meron cautiously warned, 'the Martens clause does not allow one to build castles of sand'.<sup>103</sup> Contrary to what the supporters of the pro-ban campaign seem to imply, States never meant to include the principles of humanity and the dictates of public conscience among the sources of international humanitarian law.<sup>104</sup> Still, as the same Meron pointed out, the Clause performs a more limited, but equally valuable function, in that '[i]t serves as a powerful vehicle for governments and especially NGOs to push the law ever more to reflect human rights concerns. Where there already is some legal basis for adopting a more humanitarian position, the Martens clause enables decision makers to take the extra step forward'.<sup>105</sup>

The latter observation allows us to better frame the Martens Clause within the overall pro-ban campaign's argumentative strategy, by pinpointing its precise role and weight. As they are currently understood, neither the 'principles of humanity' nor the 'dictates of public conscience' prongs of the Clause can work as self-standing arguments against AWS, as they do not constitute formal sources of international law. Rather, they may reinforce the arguments analysed in the previous sections, providing pro-ban negotiators with additional leverage to push

<sup>101</sup> Written Statement and Comments of the Russian Federation on the Issue of the Legality of the Threat or Use of Nuclear Weapons (16 June 1995) 13.

<sup>102</sup> Advisory Opinion on *Nuclear Weapons* (n 91) para 78.

<sup>103</sup> Meron (n 89) 88.

<sup>104</sup> This has been persuasively demonstrated by Cassese (n 78), in the light of the *travaux préparatoires* relating to the Clause as well as of the practice applying it.

<sup>105</sup> Meron (n 89) 88.

hesitant States to move forward towards the adoption of a Protocol prohibiting the lethal uses of AWS.

6. *Autonomy in weapons systems constitutes a threat to international peace and security*

A criticism commonly levelled against pro-ban campaigners is that they completely obliterate the positive impact that autonomy in weapons systems could have on the protection of innocent civilians, and the respect for IHL in general. AWS, it is argued, can become more accurate than human soldiers in targeting military objectives and, unlike human soldiers, are utterly unconstrained by the need for self-preservation and immune from human passions (such as anger, fear and vengefulness).<sup>106</sup>

The strength of this contention ultimately depends on a key assumption, namely that the deployment of AWS will not have a significant impact outside battlefield scenarios. Yet, this is a narrow appraisal that only captures a fraction of the overall picture, since it screens off more pervasive effects that are likely to flow from an increased use of AWS.<sup>107</sup> Reference is made, in particular, to the fact that autonomy in weapons systems could yield nefarious effects on the global level in that it would make wars easier to wage,<sup>108</sup> with heavy backlashes on international peace and security.<sup>109</sup> This enlarged perspective brings into play a dis-

<sup>106</sup> Schmitt (n 45) 23; Sassòli, *Autonomous* (n 85) 310; R Jenkins, D Purves, 'Robots and Respect: A Response to Robert Sparrow' (2016) 30 *Ethics & Intl Affairs* 391, 396-399; Anderson, Waxman (n 49) 1108. In fact, this argument is partly weakened by the abovementioned consideration that, realistically, AWS will not be able (if ever) to outperform humans in distinction and proportionality calculations for a very long time.

<sup>107</sup> Tamburrini (n 14) 137-141.

<sup>108</sup> As the Austrian delegation observed in its Opening Statement at the 2014 CCW Meeting of Experts: '[p]utting soldiers' lives at stake makes States think twice whether to engage in armed conflict. Autonomous weapons remove such restraint from the use of military force'. It is remarkable that similar concerns have been recently voiced by the Chinese delegation in a position paper submitted at the CCW Fifth Review Conference, (<[www.unog.ch/80256EDD006B8954/\(httpAssets\)/DD1551E60648CEBBC125808A005954FA/\\$file/China's+Position+Paper.pdf](http://www.unog.ch/80256EDD006B8954/(httpAssets)/DD1551E60648CEBBC125808A005954FA/$file/China's+Position+Paper.pdf)>).

<sup>109</sup> For a wider analysis of the possible destabilising effects of AWS, see International Committee for Robot Arms Control, 'LAWS: Ten Problems for Global Security'



tinct legal regime, since we move from the law regulating the conduct of hostilities (IHL, or *jus in bello*) to the rules governing the use of force, or *jus ad bellum*.

A preliminary observation on the shifting legal framework is in order here. While the potential impact of AWS on *jus ad bellum* is often described as a matter of concern for international law, this proposition has rarely been discussed in depth, so that it is not entirely clear what the actual legal issue at stake is. This led one author to rule out the relevance of *jus ad bellum* in this field, in the light of the fact that the determination as to '[w]hether a breach of a rule of *ius ad bellum* has occurred [...] is independent from the type of weapon that has been used'.<sup>110</sup> This sounds entirely reasonable. Whether a certain use of force is contrary to the *jus ad bellum* ultimately depends on the circumstances in which force is unleashed (eg whether the use of force was in self-defence or was authorised by the UN Security Council), and not on the sorts of weapons that are employed.

Yet, this is not the whole story. The law governing the maintenance of international peace and security should not be reduced to a static, binary rule. This legal regime is not only about determining whether a specific armed activity is lawful or not under the prohibition on the use of force. Rather, it is about ensuring – in the words of the 1984 Declaration on the Right of Peoples to Peace – that 'the policies of States be directed towards the elimination of the threat of war'.<sup>111</sup> This claim entails that a more comprehensive (and dynamic) appraisal must be carried out, which may well include an evaluation of policies allowing the use of AWS, especially in connection with the question whether these policies are conducive to more peace and security in international relations or, on the contrary, represent a factor of instability at global and regional levels.

In effect, a policy allowing the use of AWS would end up encouraging a more liberal approach to the use of force by States and, as a consequence, a higher likelihood of violations of the prohibition on the use of force under Article 2(4) UN Charter. It has been shown elsewhere<sup>112</sup> that the more human troops are replaced by AWS, the less will be the

(April 2015) <<http://icrac.net/wp-content/uploads/2015/04/LAWS-10-Problems-for-Global-Security.pdf>>.

<sup>110</sup> Wagner (n 58) marginal number 11.

<sup>111</sup> UNGA Resolution 39/11 (1984) para 3.

<sup>112</sup> Amoroso, Tamburrini (n 34) 10-11.

potential costs of conflicts in terms of human losses and, with it, sensitivity to the issue in the general public. This would in turn relax popular and parliamentary control over governmental exercise of war powers, both of which have sometimes proved decisive in preventing (or stopping) military actions in breach of international law.<sup>113</sup>

It could be objected that these concerns are not specific to AWS, as the latter are not the sole military technology aimed at removing soldiers from the battlefield.<sup>114</sup> One may think of armed drones, which have essentially immunised human attackers by placing them kilometres away from combat theatres.<sup>115</sup> In fact, this objection does not sufficiently consider that autonomy in weapons systems, to the extent that it replaces remote control, would let unmanned technologies make a qualitative leap forward, by overcoming the limitations that have so far precluded their large scale use in aggressive military actions: AWS cannot be jammed, they do need not huge bandwidth from satellites, nor do they have to be in constant line of communication with human operators, which makes their reactions faster and hence far more efficient in real time combat situations.<sup>116</sup>

## 7. Concluding remarks

Since the launch of the Campaign, its legal tenets have been under fire by critics, who have deployed a battery of counter-arguments to affirm the legality and ethical acceptability of the lethal uses of AWS. This paper has tried to show that, notwithstanding that several attacks have

<sup>113</sup> See, for instance, with regard to the US decision not to wage (an arguably unlawful) war against Assad in Syria, H Dieck, *The Influence of Public Opinion on Post-Cold War U.S. Military Interventions* (Palgrave MacMillan 2015) 185.

<sup>114</sup> Sparrow (n 87) 106.

<sup>115</sup> On the impact of the use of drones on democratic check and balances, especially in connection with their use in the US 2011 Libya intervention, see E Datteri and G Tamburrini, 'Robotic Weapons and Democratic Decision-Making' in E Hilgendorf, J-P Günthe (eds), *Robotik und Gesetzgebung* (Nomos 2013) 211.

<sup>116</sup> I Pejic, 'Autonomous Weapon Systems and US Military Robotics: Unmanned Aerial and Ground Vehicles (UAV and UGV)' *Global Research* (4 February 2016) <[www.globalresearch.ca/autonomous-weapon-systems-and-us-military-robotics-unmanned-aerial-and-ground-vehicles-uav-and-ugv/5505415](http://www.globalresearch.ca/autonomous-weapon-systems-and-us-military-robotics-unmanned-aerial-and-ground-vehicles-uav-and-ugv/5505415)>.



been carried out, no one attack has dealt a deadly blow. At the same time, however, it has brought into the limelight that pro-ban arguments are not equal in strength and weight, although they all perform a role in making the case against autonomy in lethal weapons systems (or, which is ultimately the same, in favour of a meaningful human control over them). Awareness of this, it is here submitted, could help the campaigners to parry more effectively the opponents' argumentative blows. To conclude, let us briefly summarise the main outputs of the foregoing analysis.

The contention whereby AWS are unable to comply with IHL principles, which constitutes the bulk of the pro-ban campaign, has resisted the various criticisms brought against it, *but only insofar as current and foreseeable AWS are concerned*. Its validity is thus temporally limited, as one cannot rule out *in abstracto* the feasibility of IHL-compliant AWS, although this will probably occur in the very distant future. It therefore needs to be reinforced by arguments that do not suffer from the same inherent limitations.

Reference to the AWS-related accountability gap can certainly serve this purpose, as it comes across as the most solid argument at disposal of the Campaign. Notwithstanding the numerous attempts to identify alternative forms of responsibility, indeed, the fact remains that – regardless of technological advancements – in most cases no one will be *directly* held accountable for war crimes committed by AWS.

On the other hand, arguments based on the Martens Clause, with particular regard to its 'public conscience' prong, perform a more limited function. As such, they are not able to justify a prohibition on Lethal AWS. Nevertheless, widespread (and qualified) aversion towards non-human lethal decision-making could be an additional push factor in the negotiation process.

Finally, *jus ad bellum* arguments play a two-fold role. On the one hand, they help neutralise the objection whereby a ban on Lethal AWS would be in fact be anti-humanitarian in that autonomy in weapons systems could ensure more accuracy in targeting and less casualties. On the other hand, as suggested elsewhere,<sup>117</sup> they bolster the case for an extensive AWS ban reaching beyond potentially lethal uses of these weapons systems, so as to also cover non-lethal AWS with expected destabilising effects on international peace and security.

<sup>117</sup> Amoroso, Tamburrini (n 34) 15.