



Università degli Studi di Napoli Federico II

**THE CONVERGENCE OF DEONTOLOGICAL
AND CONSEQUENTIALIST REASONS FOR BANNING
AUTONOMOUS WEAPONS SYSTEMS**

Daniele Amoroso

Dipartimento di Giurisprudenza

Guglielmo Tamburrini

DIETI – Dipartimento di Ingegneria Elettrica e Tecnologie
dell'Informazione

12 October 2016 (Preprint)

The convergence of deontological and consequentialist reasons for banning Autonomous Weapons Systems

Daniele Amoroso* and Guglielmo Tamburrini[^]

Abstract

This paper examines deontological and consequentialist reasons for banning autonomous weapons systems (AWS), exploring their implications for the meaningful human control of weapons. First, the main deontological and consequentialist arguments leveled against AWS from ethical and legal perspectives are examined with the aim of assessing their respective scope, strengths and limitations. Second, it is shown that these strengths and limitations are largely complementary. Third, it is argued that deontological and consequentialist arguments can be coherently combined so as to provide a set of mutually reinforcing ethical and legal reasons for banning AWS. This is achieved on the basis of a confluence model which prioritizes deontological reasons over consequentialist ones. Since deontological arguments bring out the obligations of special sorts of agents and the rights of special sorts of patients, there are circumstances in which these obligations and rights do not apply. There, the proposed confluence model applies consequentialist arguments that are neutral with respect to both agents and patients. Finally, it is argued that the proposed confluence model significantly bears on the issue of what it is to exercise meaningful human control on existing weapons systems which can already operate in autonomous mode.

1. Introduction

The morality and legality of autonomous weapons systems (AWS) is currently in the spotlight. Far from being just a topic of discussion within academic circles, this is an issue of common concern for States, International Organizations and NGOs, which have been debating the way forward on AWS, on a fairly regular basis, both in formal and informal fora.¹ Remarkably enough, since 2014, an Informal Meeting of Experts on Lethal AWS has been held annually in Geneva within the institutional framework of the Convention on Conventional Weapons (hereinafter CCW Meeting of Experts),

* Researcher in International Law, Dipartimento di Giurisprudenza, Università di Napoli Federico II, Italy (da.amoroso@gmail.com).

[^] Professor of Logic and Philosophy of Science, DIETI – Dipartimento di Ingegneria Elettrica e Tecnologie dell'Informazione, Università di Napoli Federico II, Italy (tamburrini@unina.it). While the authors equally share the responsibility for the entire work, just for evaluation purposes sections 2, 4.1 and 5 should be attributed to Guglielmo Tamburrini, while the remaining sections should be attributed to Daniele Amoroso.

¹ A fairly comprehensive chronology of the initiatives taken in this respect at the international and national level is available at <http://www.stopkillerrobots.org/chronology/>.

gathering an ever-increasing number of States.² Awareness about this topic has been significantly raised by the Campaign “Stop Killer Robots”, which was launched in 2013 by an international coalition of NGOs with the primary goal of “ban[ning] lethal robot weapons”.³

This campaign has received wide (although far from unanimous) support in the academia.⁴ Scholars favoring a ban on AWS come from different fields of inquiry (robotics, military science, philosophy, international law). However, the gist of their argumentation lies at the intersection between normative ethics and international law.

In normative ethics, reasons for banning AWS have been put forward from both deontological and consequentialist standpoints. Roughly speaking, deontological arguments aim to support the conclusion that by deploying AWS one is likely or even bound to violate moral obligations of special sorts of agents (military commanders and operators) or moral rights of special sorts of patients (AWS potential victims). Consequentialist arguments support the conclusion that the choice of prohibiting AWS is expected to protect peace and security, thereby enhancing collective human welfare, and more effectively so than the incompatible choice of permitting their use.

Arguments based on international law have been developing along similar lines. On the one hand, it has been maintained that the deployment of AWS would be in stark contrast with basic tenets of international humanitarian law (IHL), international human rights law (IHRL) and international criminal law (ICL). These legal arguments are strictly related to reasons advanced from an ethical perspective insofar as both the obligations of military operators and the rights of potential AWS victims enshrined in IHL, IHRL and ICL are morally grounded in deontological ethical theories. On the other hand, in an ostensibly consequentialist perspective, the deployment of AWS has been claimed to make wars easier to wage, with inevitable backlashes on the prohibition on the use of force (art. 2(4) UN Charter) and, more generally, on the maintenance of international peace and security (art. 1(1) UN Charter).

Both argumentative lines have been criticized by a number of scholars, skeptical of the ban, who emphasized their purported weaknesses and inherent limits. While ultimately supporting the reasons for a ban, the present paper acknowledges some limitations of each pro-ban argument taken in isolation. Nevertheless, it is argued here that, their respective limitations notwithstanding, deontological and consequentialist arguments strongly support the case for a ban on AWS, because of their

² All the transcripts and documents of the CCW Meetings of Experts are available at [http://www.unog.ch/80256EE600585943/\(httpPages\)/8FA3C2562A60FF81C1257CE600393DF6?OpenDocument](http://www.unog.ch/80256EE600585943/(httpPages)/8FA3C2562A60FF81C1257CE600393DF6?OpenDocument). A formal discussion on AWS among CCW State Parties will be held at the Fifth Review Conference of the Convention on Conventional Weapons (CCW) (Geneva, 12-16 December 2016).

³ See the launch statement at http://www.stopkillerrobots.org/wp-content/uploads/2013/03/KRC_LaunchStatement_23Apr2013_fnl.pdf.

⁴ The call for a ban has been also endorsed by the EU Parliament, which adopted a resolution 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” (European Parliament resolution of 27 February 2014 on the use of armed drones (2014/2567(RSP)), para. 2).

complementary and mutually reinforcing character. Accordingly, a confluence model for pro-ban deontological and consequentialist approaches is advanced, which enables one to solve potential conflicts between these two approaches. Moreover, it is maintained that these arguments significantly bear on the issue of what meaningful human control is to be exercised on existing AWS and, relatedly, on the definition of the scope of a future AWS ban treaty.

Preliminary to this analysis, various requirements that have been put forward for a weapons system to count as autonomous are analyzed in section 2. As a result of this analysis, the methodological choice is made of adopting the more liberal requirements, which are satisfied by various existing weapon systems, instead of the more restrictive requirements which end up projecting AWS in some undetermined technological future, and defer the need for regulating their production and use accordingly.

Deontological perspectives on banning AWS are examined in section 3. More specifically, three clusters of arguments are discerned on account of which conclusion they are intended to support. First, there is the conclusion that AWS are unable to comply with the rules governing the use of lethal force in both IHL and IHRL (section 3.1). Second, to the extent that they take humans “out-of-the-loop”, it is argued that AWS are likely to determine an accountability gap, which is hardly reconcilable with the principle of individual criminal responsibility under ICL (3.2). Third, it is maintained that the deployment of AWS would run contrary human dignity, which is a foundational value of both IHL and IHRL, and which commands that the taking of human life should be reserved to human decision-makers (3.3). Arguments for the first conclusion are found to depend on technological limitations affecting current and foreseeable AWS. As a consequence, these arguments are found to effectively support limited actions only, towards the establishment of a moratorium or the development of special regulation on AWS permissible uses. In contrast with this, arguments for the remaining two claims are independent of technologically contingent situations, and make rather appeal to ethically and legally entrenched obligations of special sorts of agents or to some patient-relative human rights. Accordingly, these arguments are found to effectively support a prohibition of AWS *lethal* uses (3.4).

Consequentialist perspectives on AWS are scrutinized in section 4. There, it is maintained that a wide appraisal of expected AWS deployment consequences (as opposed to narrow consequentialist appraisals that are confined to battlefield scenarios and performances) is conducive to the conclusion that various prospective AWS are likely to bring about significant destabilizing effects, thereby raising ethical (4.1) and legal concerns (4.2) about peace and stability preservation at both regional and global scales. The scope of this wide consequentialist approach to AWS production and use is investigated too. In particular, this approach is found to provide a solid motivation for developing legal solutions enabling one to regiment AWS with significant destabilizing effects on peace and security (4.3).

The convergence of deontological and consequentialist perspectives offers reasons to prohibit a wide range of both lethal and non-lethal uses of AWS (their non-lethal uses including, e.g., hostilities conducted by means of opposing teams of AWS and without the involvement of human soldiers). Accordingly, in order to assess the overall strength of ethical and legal arguments in support of an AWS ban that have been variously advanced from deontological and consequentialist perspectives, it is crucial to understand whether deontological and consequentialist reasons can be coherently combined into a unitary ethical and legal framework. In section 5, a confluence model is suggested, which prioritizes deontological perspectives whenever agent-relative obligations and patient-relative rights are applicable, and shifts by default to agent-neutral consequentialist perspectives whenever those obligations and rights are no longer applicable. Finally, it is demonstrated that this confluence model can be informatively applied to determine what it is to exert meaningful human control on existing weapon systems which, according to the more liberal requirements for autonomy, can already operate in autonomous mode.

2. On the notion of autonomy in AWS ethical and legal debates

Ethical and legal debates on AWS are usually based on a construal of autonomy as some kind of task-execution or goal-achievement capabilities and related information-processing skills for technological devices.⁵ A demanding definition of AWS along these lines was advanced in a document of the UK Ministry of Defense. There, an autonomous system is required to be “a system which is capable of *understanding higher level intent and direction*. From this understanding and its perception of its environment, such a system is able to take appropriate action to bring about the desired state”.⁶ This requirement is hardly satisfied by any existing weapons system. Moreover, no educated guess can be made about the prospects of constructing machines which are capable of understanding high level intent and direction, unless one redefines these terms operationally and quite differently from common usage. Therefore, on the basis of common linguistic usage, the UK Ministry of Defense definition projects AWS in some undetermined technological future.⁷

The US Department of Defense (DoD) proposed a less demanding requirement on weapons systems to count as autonomous. Indeed, autonomous are those weapons systems ‘that, once activated, can

⁵ 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), p. 122 ff., pp. 124-127.

⁶ UK Ministry of Defence, *The UK Approach to Unmanned Aircraft Systems*, Joint Doctrine Note 2/11, 30 March 2011, p. 14, emphasis added (https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/33711/20110505JDN_211_UAS_v2U.pdf).

⁷ Coherently with this premise, the UK has boldly affirmed that AWS “do not, and may never, exist”. See the Opening Statement by the UK delegation at the 2016 CCW Meeting of Experts (General Exchange, p. 1).

select and engage targets without further intervention by a human operator.”⁸ Much of the ongoing ethical, legal and political discussion on AWS takes the DoD requirement for autonomy as a starting point.⁹ Nevertheless, in these discussions the DoD requirement is usually refined and suitably adapted to various classes of AWS, insofar as the required autonomy is always relative to some set of warfare environments, and to the specific perceptual, reasoning and action capabilities that the weapon system is endowed with.¹⁰

One should carefully note that the DoD requirement licenses as autonomous a variety of weapons systems that are already operating. These include defensive systems like the US Aegis and the Israeli Iron Dome, a variety of offensive fire-and-forget munitions, like the British Brimstone missile, and stationary (i.e. non-mobile) robotic sentinels, like the South-Korean SGR-A1, which surveils the border between North and South Korea, or the gun tower developed by the Turkish company ASELSAN to be installed on the border between Turkey and Syria.¹¹ Thus, if one accepts the DoD requirement as an adequate criterion to discriminate between autonomous and non-autonomous weapons systems, then one is *ipso facto* acknowledging that some AWS exist and have been actually deployed in various warfare scenarios.¹² Accordingly, the ethical, legal, and political discussions of AWS that are based on the DoD requirement cannot be construed as discussions concerning future combat systems only. This far-reaching implication of the methodological choice of accepting the DoD requirement in ethical, legal, and political discussions of AWS is aptly illustrated by reference to Iron Dome, Brimstone, and SGR-A1. Let us briefly consider each one of these systems in turn.

The Iron Dome system was primarily designed to counter incoming short-range rockets. Human operators feed the system in advance with information concerning the area to be protected from incoming rockets. On the basis of this information, the system detects the launch of a rocket and tracks its trajectory in order to determine whether the incoming projectile threatens the designated area. Only

⁸ US Department of Defense, *Autonomy in Weapons Systems*, Directive 3000.09, 21 November 2012, pp. 13–14 (available at www.dtic.mil/whs/directives/corres/pdf/300009p.pdf).

⁹ See, among many others, Human Rights Watch, *Losing Humanity. The Case against Killer Robots* (HRW 2012), pp. 7-8 ; C. Heyns, “Report of the Special Rapporteur on extrajudicial, summary or arbitrary executions”, 9 April 2013, UN Doc. A/HRC/23/47 (hereinafter, “Heyns Report”), para. 38; M.P. Schmitt, “Autonomous Weapon Systems and International Humanitarian Law: A Reply to the Critics”, *Harvard National Security Journal Features* (2013), p. 4; R. Crootof, “The Killer Robots Are Here: Legal and Policy Implications”, *Cardozo Law Review* (2015), p. 1837 ff., p. 1847; E. Liebllich and E. Benvenisti, “The obligation to exercise discretion in warfare: why autonomous weapons systems are unlawful”, in N. Bhuta et al. (eds.), *cit. supra* note 5, p. 245 ff., p. 249; D.P. Hollis, “Setting the Stage: Autonomous Legal Reasoning in International Humanitarian Law”, *Temple International & Comparative Law Journal* (forthcoming, available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2711304), p. 6.

¹⁰ See, for instance, the Working Paper presented by International Committee of the Red Cross at the 2016 CCW Meeting of Experts (“Views of the International Committee of the Red Cross (ICRC) on autonomous weapon system”, 11 April 2016, pp. 3-4).

¹¹ “Turkey starts building automatic shooting gun towers at Syrian border”, *Yeni Şafak*, 30 May 2016, available at <http://www.yenisafak.com/en/news/turkey-starts-building-automatic-shooting-gun-towers-at-syrian-border-2473487>. For more information about the ASELSAN Stabilized Gun Platform, see <http://www.aselsan.com.tr/en-us/capabilities/weapon-systems>.

¹² R. Crootof, “The Killer Robots Are Here”, *cit. supra* note 9. See also the ICRC’s Working Paper, *cit. supra* note 10, p. 2, and the Statement by Sierra Leone at the 2016 CCW Meeting of Experts (General Exchange, p. 2).

if a threat is detected, an interceptor missile is fired to destroy the incoming rocket.¹³ Accordingly, the critical target selection and attack functionalities mentioned in the DoD requirement are both performed without any human intervention. A naval version of Iron Dome (called C-Dome) – which is being developed and tested to protect vessels from missile, UAV and other types of attack weapons – can similarly perform these critical functionalities without any human intervention.¹⁴

Brimstone is an air-to-ground attack missile which works either as fire-and-forget munition or as laser-guided missile. In the fire-and-forget mode, Brimstone is loaded with targeting data, including data which serve to circumscribe the area within which armored enemy vehicles must be searched for, selected and attacked without any further intervention by human operators. In the laser-guided mode, human operators pick out the specific targets to destroy.¹⁵ Thus, Brimstone missiles satisfy the DoD requirement for autonomy in the fire-and-forget mode, and fail to satisfy the same requirement in the laser-guided mode. The operation mode is selected by human operators on the basis of available information about individual attack scenarios (e.g. considering whether there are civilians or friendly forces in the vicinity of targets).

The Samsung system SGR-A1 is a stationary robotic platform designed to replace or to assist South Korean sentinels in the surveillance of the demilitarized zone between North and South Korea. The SGR-A1 can be operated in either unsupervised or supervised modes. In the unsupervised mode, the SGR-A1 identifies and tracks intruders in the demilitarized zone, deciding whether to fire at them or not without any further intervention by human operators. In the supervised mode, firing actions are contingent on the judgment and the “go” command of a human operator.¹⁶ Thus, the SGR-A1 satisfies the DoD requirement if it operates in the unsupervised mode, and it does not otherwise. In the latter case, it is best viewed as a combination of a decision-support system with a remote-controlled firing device. It is worth noting that the supervised SGR-A1 preserves most of the critical functions in the DoD requirement, insofar as it performs without any human intervention the tasks of target identification and tracking in its intended operational environment. Accordingly, this robotic sentinel affords a straightforward illustration of the fact that a simple on/off operational mode switch can make all the difference between an autonomous and a non-autonomous weapon system on the basis of the DoD requirement on AWS as a binary discrimination criterion.

By accepting the DoD requirements as a binary discrimination criterion, one is *ipso facto* committed to recognize the existence and actual deployment of several types of AWS, contrary to what was asserted by various representatives of State parties at CCW informal meetings of experts in Geneva between

¹³ See <http://www.army-technology.com/projects/irondomeairdefencemi/>.

¹⁴ T. Eshel, “Rafael extends Iron Dome C-RAM to the naval domain”, *Defense Update*, 26 October 2014, available at http://defense-update.com/20141026_c-dome.html#.VE_lQ8IcQdU.

¹⁵ See <http://www.army-technology.com/projects/brimstone/>

¹⁶ G. Tamburrini, *cit. supra* note 5, p. 126.

2014 and 2016.¹⁷ Moreover, it is worth noting that the distinctions between stationary/mobile or defensive/offensive weapons systems are irrelevant to the problem of deciding whether a weapons system should be counted as autonomous or not on the basis of the DoD criterion: SGR-A1 is a stationary system, and the Brimstone missile is not; the former is a defensive system and the latter is not. Fine-grained distinctions in terms of the types and levels perceptual and cognitive capabilities appear to be similarly immaterial with respect to this discrimination problem, insofar as the pattern matching capabilities that are required as a precondition for attacking a short-range rocket are quite different from the perceptual and cognitive capabilities that are required of a robotic sentinel to recognize a human target or of an air-to-ground missile to identify armored enemy vehicles chiefly on the basis of their visual models.

Using the DoD requirements as an operational criterion enabling one to distinguish between autonomous and non-autonomous weapons systems is in some respects unsatisfactory, at least insofar as these requirements force one to place in the same AWS class extant defensive systems like the Iron Dome, which exclusively reacts to and targets incoming short range rockets, and any future offensive system which will be supposedly capable to operate autonomously in the fog of war, and to discriminate active foes from civilians and *hors de combat* enemies in cluttered urban warfare scenarios. It was noted above that this inconvenience cannot be sensibly overcome by adopting the prohibitively restrictive UK definition of AWS. At the same time, in order to introduce a new definition of AWS striking a good balance between the DoD and the UK definitions, one has to address the formidable problem of identifying a non-arbitrary cutting point in between the perceptual, cognitive, and action capabilities of systems like the Iron Dome on the one hand, and those characterizing the envisaged AWS that will be capable of matching infantryman skills in cluttered urban warfare scenarios.¹⁸

In view of these interrelated difficulties, and without further ado, the DoD requirement is adopted here too as a basis for distinguishing between AWS and other weapons systems, postponing a discussion about the need for more refined distinctions between different types of AWS after an examination of the main arguments which have been so far advanced against AWS. To begin with, let us consider deontological arguments.

¹⁷ In addition to the stance taken by the UK, mentioned above (see *supra* note 7), this view was expressed, on the basis of equally strict notions/requirements of autonomy, by Argentina (2014 CCW Meeting of Experts, Closing Statement), Denmark (2015 CCW Meeting of Experts, General Exchange), France (2016 CCW Meeting of Experts, Working Paper on “[Characterization of a Lethal AWS](#)”, p. 2), Germany (2015 CCW Meeting of Experts, Final Statement, p. 2), Italy (2016 CCW Meeting of Experts, Statement at the Panel “Towards a Working Definition of Lethal AWS”, p. 2), Japan (2016 CCW Meeting of Experts, Working Paper, p. 2), Sweden (2014 CCW Meeting of Experts, Opening Statement, p. 1), and Turkey (2016 CCW Meeting of Experts, General Exchange).

¹⁸ It is worth noting in passing that a similar difficulty has long afflicted discussions of what is intelligence and intelligent behavior in connection with the aims and achievements of Artificial Intelligence (AI), Cognitive Robotics and other cognate disciplines.

3. Deontological perspectives on banning AWS

A rich family of arguments for banning AWS are framed in deontological terms, with solid underpinnings in international law. Deontological arguments for banning AWS can be informatively classified into three groups on the basis of the claims that each one of them is meant to buttress. First, AWS would be unable to comply with various obligations flowing from both IHL and IHRL rules governing the use of lethal force. Second, AWS, to the extent that they take humans “out-of-the-loop”, are likely to determine an accountability gap. And the latter is hardly reconcilable with the agent-relative moral obligation of military commanders and operators to be accountable for their own actions, as well as with the related principle of individual criminal responsibility under ICL. Third, the deployment of AWS would run contrary to a central foundational value of both IHL and IHRL, namely human dignity and the patient-relative rights that come with it, which require that the taking of human life should be reserved to human decision-makers.

Let us turn to examine the main reasons that have been adduced for each one of these claims.

3.1. AWS are unable to comply with IHL and IHRL rules governing the use of lethal force

The main legal norms appealed to in deontological arguments for banning AWS are the IHL rules of distinction¹⁹ and proportionality.²⁰ It is seriously doubted that AWS will be able, in the foreseeable future, to comply with these rules in a wide variety of warfare scenarios.²¹ An AWS complying with IHL requirements is usually taken to be an autonomous weapon which is capable of respecting the

¹⁹ Art. 48 of the First Additional Protocol to the Geneva Conventions (AP I) (“[...] the Parties to the conflict shall at all times distinguish between the civilian population and combatants and between civilian objects and military objectives and accordingly shall direct their operations only against military objectives”). See also Arts. 51(2) (“The civilian population as such, as well as individual civilians, shall not be the object of attack. [...]”) and 52(1) (“Civilian objects shall not be the object of attack or of reprisals. [...]”). This rule has acquired the status of customary international law (J.-M. Henckaerts and L. Doswald-Beck, *Customary International Humanitarian Law* (ICRC 2006), pp. 3-8).

²⁰ According to the codification of customary IHL rules edited by the ICRC, the rule of proportionality prohibits “[l]aunching an attack which may be expected to cause incidental loss of civilian life, injury to civilians, damage to civilian objects, or a combination thereof, which would be excessive in relation to the concrete and direct military advantage anticipated, is prohibited” (J.-M. Henckaerts and L. Doswald-Beck, *ibid.*, p. 46). References to this rule may be found in Arts. 51(5)(b) and 57(2)(a)(iii) and (b) AP I.

²¹ N.E. Sharkey, “The inevitability of autonomous robot warfare”, *International Review of the Red Cross* (2012), p. 787 ff.; D. Akerson, “The Illegality of Offensive Lethal Autonomy”, in D. Saxon (ed.), *International Humanitarian Law and the Changing Technology of War* (Martinus Nijhoff, 2013), p. 65 ff.; Heyns Report, paras. 63-74; C. Grut, “The Challenge of Autonomous Lethal Robotics to International Humanitarian Law”, *Journal of Conflict & Security Law* (2013), p. 5 ff.; K. Egeland, “Lethal Autonomous Weapon Systems under International Humanitarian Law”, *Nordic Journal of International Law* (2016), p. 89 ff. This issue has also been raised, although with different overtones, by many of the delegations intervened at the CCW Meetings of Experts. See the views expressed by Argentina (2014 CCW Meeting of Experts, Closing Statement), Austria (2015 CCW Meeting of Experts, Working Paper “The concept of ‘meaningful human control’”), Cuba (2015 CCW Meeting of Experts, Working Paper, p. 1), Ecuador (2016 CCW Meeting of Experts, General Exchange, pp. 2-3), the Holy See (2015 CCW Meeting of Experts, Statement at the Panel on “Overarching Issues”, p. 3), Norway (2014 CCW Meeting of Experts, Opening Statement, pp. 1-2); Pakistan (2015 CCW Meeting of Experts, General Exchange, p. 4), South Africa (2014 CCW Meeting of Experts, Opening Statement, p. 2), Sweden (2014 CCW Meeting of Experts, Statement at the Panel on “Legal Aspects – International Humanitarian Law”), Switzerland (2016 CCW Meeting of Experts, Statement at the Panel on “Challenges to International Humanitarian Law”, p. 3), United Kingdom (2015 CCW Meeting of Experts, Statement at the Panel on “Challenges to International Humanitarian Law”, p. 2).

principles of distinction and proportionality *at least as well as a competent and conscientious human soldier*. The possibility of an AWS fulfilling IHL in this sense presupposes the solution of many profound research problems in advanced robotics and AI.²²

On the one hand, while present-day UAVs “cannot distinguish a sleeping dog from a bush, even at high noon,”²³ the rule of distinction raises the problem of endowing AWS with the capability of perceptually distinguishing civilians and *hors de combat* people in erratic and surprise seeking warfare environments, in addition to the capability of establishing whether civilians have lost protection from attacks because of their participation into hostilities.²⁴ Thus, in particular, AWS must be capable of recognizing behaviors that conventionally or unconventionally carry hostile or surrender messages and fighting incapacitation information, on the basis of a variety of perceptual and cognitive capabilities, including viewpoint-independent recognition of bodily postures and gestures in the variable perceptual conditions of unstructured warfare scenarios, understanding of emotional expressions, and real-time reasoning about deceptive intentions and actions.

On the other hand, the proportionality rule requires the decision-maker to strike a delicate balance between military gains expectedly deriving from some given course of action and harms to civilians (in terms of life loss and injuries or damages to civilian objects) ensuing from it. In fact, it is argued, a proper appraisal of the terms of this balance, as well as of their interactions can hardly be carried out by present and foreseeable AI systems,²⁵ especially insofar as it requires advanced social interaction competence and high sensitivity to the emotional and cognitive context, in addition to the military commander’s evaluation of human factors which include the shared viewpoints and feelings arising in soldiers on account of extended periods of communal daily life, training, and field experience.

This point can be further elucidated with an example. Let us consider a simplified military scenario where the decision-maker faces the following two options: Option A, which will yield a substantial military advantage, is likely to provoke a significant amount of civilian casualties; and Option B, which will cause less damage to the enemy, is unlikely to harm civilians. A military commander who has decided that the proportionality principle is respected in option A might still refrain from following this course of action on account of the nefarious impact which, knowing well her or his soldiers and sharing with them daily warfare hardships, she or he deems that the killing of civilians will have on the troops’

²² N.E. Sharkey, ‘Saying ‘no!’ to lethal autonomous targeting’, *Journal of Military Ethics*, 9 (2010), p. 369 ff., p. 378; R. C. Arkin, *Lethal autonomous systems and the plight of the non-combatant*, *AISB Quarterly* 137, pp. 1-9, p. 4.

²³ L. G. Weiss, “Autonomous Robots in the Fog of War” *IEEE Spectrum* (2011), <http://spectrum.ieee.org/robotics/military-robots/autonomous-robots-in-the-fog-of-war/>

²⁴ P. Alston, “Lethal Robotic Technologies: The Implications for Human Rights and International Humanitarian Law”, *Journal of Law Information and Science* (2011), p. 35 ff., p. 54; N.E. Sharkey, *cit. supra* note 21, pp. 788-789; Grut, *cit. supra* note 21, pp. 11-12; R. Geiss, “The International-Law Dimension of Autonomous Weapons Systems” (Friedrich-Ebert-Stiftung, 2015), p. 14.

²⁵ N.E. Sharkey, *cit. supra* note 21, pp. 789-790; Heyns Report, paras. 70-73; Grut, *cit. supra* note 21, pp. 12-14; Egeland, *cit. supra* note 21, pp. 103-105.

morale. This negative effect will in fact, according to her or his best judgment, jeopardize in the long run the possibility of defeating the enemy in the given operational theatre. It is presently difficult to imagine technologically foreseeable machines which are capable of gauging the psychological state of the troops involved in a conflict and the related knock-on strategic effects. Accordingly, a military commander delegating in similar scenarios the task of appraising the implications of the proportionality rule and the ensuing operational decisions to an AWS would *ipso facto* violate an agent-relative obligation attached to her special role of commander in a warfare theatre.

Additional concerns have been raised in relation to the AWS' ability to comply with the rule of precaution in attacks, which supplements the aforementioned rules of distinction and proportionality by obligating warring parties to take all feasible steps to avoid, or at least to minimize, the harm caused to innocent civilians.²⁶ The focus, here, is basically on the concept of "feasibility" which, similarly to those of "civilians taking part in the hostilities" and "disproportionate collateral damage" referred above, requires complex contextual evaluations which are unlikely to be adequately performed in the foreseeable future by a non-human decision-maker.²⁷

Similar arguments have been put forward with reference to IHRL, having particular regard to the right to life and to bodily integrity. Notably, it has been submitted that, should AWS be deployed outside the context of an armed conflict, as a means of law enforcement (e.g. to quell a riot or to control a border²⁸), the thorny problems emphasized in relation to IHL would assume an even greater magnitude. Since the notions of "legitimate object of attack" and "collateral damages" are not contemplated under IHRL, lawful killings are considered there as an exception, subject to very strict conditions. Indeed, the requirements of necessity and proportionality, despite similarities in wording, are construed far more stringently than in IHL, which makes the hypothetical tasks of a law enforcement AWS even more difficult to fulfill.²⁹

3.2. Recourse to AWS determines an accountability gap

Even the more convinced proponents of the ethical acceptability and legality of AWS are compelled to admit that, no matter how accurate, these systems are error-prone in ways which are "completely alien to humans".³⁰ As a consequence, it is quite possible for an AWS to commit acts amounting – at least

²⁶ Art. 57, AP I. On the customary nature of this principle, see J.-M. Henckaerts and L. Doswald-Beck, *cit. supra* note 19, pp. 51-67.

²⁷ N. Weizmann and M. Costas Trascasas, *Autonomous Weapon Systems under International Law* (Geneva Academy of International Humanitarian Law and Human Rights, 2014), p. 16; P. Alston, *cit. supra* note 24, p. 54.

²⁸ Like the sentinel robotic guns mentioned above (see *supra* section 2).

²⁹ T. Hattan, "Lethal Autonomous Robots: Are They Legal under International Human Rights and Humanitarian Law?", *Nebraska Law Review* (2015), p. 1035 ff., pp. 1053-1054; C. Heyns, "Human Rights and the use of Autonomous Weapons Systems (AWS) During Domestic Law Enforcement", *Human Rights Quarterly* (2016), p. 350 ff., pp. 362-366.

³⁰ P. Scharre, *Autonomous Weapons and Operational Risk* (Center for a New American Security, 2016), p. 17.

materially – to war crimes or crimes against humanity.³¹ But then who will be personally responsible for these conducts? This question lies at the core of the second pro-ban argument, whereby *the delegation of the use of force to non-human decision-makers would create an accountability gap*. This proposition stands in need of some further clarification. To begin with, it goes without saying that, because of the lack of moral agency, AWS could not be held responsible as direct perpetrators.³² Responsibility for their actions, therefore, should be traced back to some persons in the decision-making chain. And that is where the real problems begin.

At the outset, one should take note that the list of potentially responsible individuals is quite long, as it includes “the software programmer, [the producer of the AWS], 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”.³³ 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,³⁴ where a group of people can be held collectively responsible for a determined outcome, whereas none of them can often be individually blamed for it.³⁵

To illustrate the many hands problem in the specific case of a war crime charge, consider the scenario of an autonomous UAV model which is programmed to distinguish between military and civilian buildings. Evidence collected in the UAV testing phase suggests that the system fails to identify as civilian objects buildings presenting certain perceptual features in 5 % of cases. In order to encourage governmental purchase, however, the manufacturer falsifies the test results, by lowering the failure rate to 0,5 %. When delivering the purchased system to the battlefield unit, the Secretary of Defense omits to mention the failure rate, but limits herself to recommending the greatest caution in the use of the aircraft against targets in the proximity of civilian objects. Basing her judgment on a liberal interpretation of this directive, a military commander deploys the aircraft over a sparsely populated area. As a consequence of this chain of events, a hospital is inadvertently bombed and several civilian victims are counted.

³¹ R. Crootof, “War Torts: Accountability for Autonomous Weapons”, *University of Pennsylvania Law Review* (2016), p. 1347 ff., pp. 1375-1377.

³² R. Sparrow, “Killer Robots”, *Journal of Applied Philosophy* (2007), p. 62 ff., p. 71; N.E., Sharkey, *cit. supra* note 21, p. 790; K. Egeland, *cit. supra* note 21, p. 91. For a contrary (but far from convincing) view, see S. Deva, “Can Robots have Human Rights Obligations? A Futuristic Exploration”, in A.S. Muller et al. (eds.), *The law of the future and the future of the law* (Torkel Opsahl Academic, 2011), vol. II, p. 185 ff.

³³ M. Wagner, “Autonomous Weapon System”, *Max Planck Encyclopedia of Public International Law* (OUP 2016), para. 21.

³⁴ H. Nissenbaum, “Accountability in a computerized society”, *Science and Engineering Ethics* (1996), p. 25 ff.

³⁵ On the many hands problem, see generally D.F. Thompson, “Moral Responsibility of Public Officials: The Problem of Many Hands”, *The American Political Science Review* (1980), p. 905 ff. This notion has been recently applied to international legal responsibility by A. Nollkaemper, “The Problem of Many Hands in International Law”, in A. Fabbricotti (ed.), *The Political Economy of International Law. A European Perspective* (Elgar, 2016), p. 278 ff.

Each character mentioned in this story contributed in distinctive ways to the accident occurrence: clearly, their acts are significant causal antecedents of the bombing event; more significantly, the conduct of each involved individual is reprehensible for some reason: the programmer should not have released the software in the face of a significant failure rate; the manufacturer should not have falsified the test results; the Secretary of Defense should have been clearer as to the risks associated with use of the AWS; the military commander should have been more cautious. And yet, none of them, taken individually, can be deemed responsible as such (both directly and indirectly) for the bombing of the hospital. Accordingly, it would be unfair to charge anyone of them with a war crime.

Accountability gaps may also occur beyond the (somewhat convoluted) “many hands” scenario. Clearly, there may be uncontroversial cases such as that of a machine which was deliberately pre-programmed to carry out international crimes or that of a commander who deployed an AWS in a context different from the one it was designed for, and where it was likely to commit war crimes.³⁶ In the majority of conceivable cases, however, the complexities of AWS technologies and their behavioral unpredictability in partially structured or unstructured warfare scenarios are likely to afford a powerful defense against criminal prosecution. Indeed, since AWS might be capable of taking courses of action whose reason “may be opaque even to the system’s designers”,³⁷ there will be cases where it is impossible to ascertain the existence of the mental element (intent, knowledge or recklessness³⁸) which is required under ICL to ascribe criminal responsibilities. As a consequence, no one could be held criminally liable, notwithstanding the conduct at stake clearly amounts to a war crime or a crime against humanity.³⁹

The doctrine of “command responsibility” is not particularly helpful either in order to fill in this accountability gap. An essential element of this doctrine is the commander’s effective control over the subordinate’s behavior. This element can be hardly applied here. On the one hand, the very notion of “autonomy” implies the absence of (or, at least, a limited role for) human supervision. On the other hand, AWS’ faster-than-human reaction times would make commander’s control to a large extent a purely speculative option.⁴⁰ Military commanders, in other words, would be stripped of the agent-relative moral duties that are traditionally associated with their rank and roles, namely, the duty to assess the risk of violating IHL requirements and to set out their orders accordingly.⁴¹

³⁶ R. Crootof, “War Torts” *cit. supra* note 31, p. 1377 (labeling them as the “easy cases”).

³⁷ *Ibid.*, p. 1373.

³⁸ It should be noted that it is fairly controversial whether “recklessness” constitutes a culpable state of mind under customary international law. See A. Whiting, “Recklessness, War Crimes, and the Kunduz Hospital Bombing”, *Just Security* (2 May 2016), available at <https://www.justsecurity.org/30871/recklessness-war-crimes-kunduz-hospital-bombing/>.

³⁹ R. Sparrow, *cit. supra* note 32, pp. 69-70; Grut, *cit. supra* note 21, pp. 14-17; Human Rights Watch, *Mind the Gap. The Lack of Accountability for Killer Robots* (HRW, 2015), pp. 19-20; R. Geiss, *cit. supra* note 24, pp. 20-21; K. Egeland, *cit. supra* note 21, pp. 110-112; R. Crootof, “War Torts” *cit. supra* note 31, pp. 1375-1378.

⁴⁰ See, also for further reference, R. Crootof, *ibid.*, p. 1380.

⁴¹ G. Tamburrini, *cit. supra* note 5, pp 131-132.

In sum, the deployment of AWS in warfare (and, possibly, law enforcement) scenarios gives rise to the serious possibility that international crimes are perpetrated, but no individual is legally responsible for those outcomes. This is in stark contrast with a “cornerstone” of contemporary international law,⁴² the principle of individual criminal responsibility.⁴³ The importance of this principle cannot be overestimated. Contrary to what some authors seem to suggest,⁴⁴ 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 the law of State responsibility. As famously noted by the Nuremberg Tribunal, in fact, 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”.⁴⁵

3.3. AWS’ lethal decision-making runs contrary to human dignity

The upshot of the third pro-ban argument is that the delegation of lethal decision-making to a machine would be prohibited under the principle of human dignity⁴⁶ (or, somehow relatedly, under the

⁴² C. Tomuschat, “The Legacy of Nuremberg”, *Journal of International Criminal Justice* (2006), p. 830 ff., p. 840

⁴³ A considerable number of States have raised this issue at the CCW Meetings of Experts. See the views set forth by Argentina (2015 CCW Meeting of Experts, General Exchange), Austria (2014 CCW Meeting of Experts, Opening Statement, p.2), Chile (2016 CCW Meeting of Experts, Statement at the Panel on “Challenges to International Humanitarian Law”, pp. 1-2), Cuba (2015 CCW Meeting of Experts, Statement at the Panel on “Overarching Issues”, p. 2), Ecuador (2016 CCW Meeting of Experts, General Exchange, p. 3), France (2015 CCW Meeting of Experts, Statement at the Panel on “Overarching Issues”, p. 2), the Holy See (2015 CCW Meeting of Experts, Statement at the Panel on “Overarching Issues”, pp. 6-7, 9), Mexico (2016 CCW Meeting of Experts, General Exchange, p. 2), Norway (2016 CCW Meeting of Experts, Statement at the Panel on “Challenges to International Humanitarian Law”, p. 2), Pakistan (2014 CCW Meeting of Experts, Opening Statement, p. 3), Poland (2015 CCW Meeting of Experts, Statement at the Panel on “Characteristics of the Lethal AWS”), Spain (2015 CCW Meeting of Experts, General Exchange, p. 2), Sri Lanka (2015 CCW Meeting of Experts, General Exchange, p. 2), Switzerland (2015 CCW Meeting of Experts, Statement at the Panel on “Challenges to International Humanitarian Law”, p. 2),

⁴⁴ R. Crootof, “War Torts”, *cit. supra* note 31, pp. 1386-1402.

⁴⁵ International Military Tribunal, Judgment of 1 October 1946, in *The Trial of German Major War Criminals. Proceedings of the International Military Tribunal sitting at Nuremberg, Germany*, Part 22 (22nd August, 1946 to 1st October, 1946), p. 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”, p. 2).

⁴⁶ At the CCW Meetings of Experts, this argument was put forth with particular strength by Germany (2014 CCW Meeting of Experts, Opening Statement, pp. 1-2) and the Holy See (2015 CCW Meeting of Experts, Statement at the Panel on “Overarching Issues”, pp. 5, 9). See also the stance taken by Chile (2015 CCW Meeting of Experts, General Exchange, p. 1), Costa Rica (2016 CCW Meeting of Experts, Statement at the Panel on “Human Rights and Ethical Issues”, pp. 1-2), Croatia (2014 CCW Meeting of Experts, Closing Statement), Cuba (2015 CCW Meeting of Experts, Statement at the Panel on “Overarching Issues”, p. 2), Ecuador (2014 CCW Meeting of Experts, Opening Statement, pp. 2-3), France (2015 CCW Meeting of Experts, General Exchange, p. 2), Morocco (2016 CCW Meeting of Experts, General Exchange, p. 3), Mexico (2015 CCW Meeting of Experts, Statement at the Panel on “Transparency and the Way Ahead”, p. 1), Pakistan (2015 CCW Meeting of Experts, Opening Statement, p. 2), Sierra Leone (2016 CCW Meeting of Expert, General Exchange, p. 3), Sweden (2015 CCW Meeting of Expert, General Exchange, p. 1), and Switzerland (2014 CCW Meeting of Experts, Opening Statement, p. 2).

“principles of humanity” enshrined in the so-called Martens Clause).⁴⁷ This argument is more far-reaching than the previous ones, as it leverages on the principle of human dignity, which is both foundational and open-textured.⁴⁸ This argument has been described as being “at the core of the concerns raised about fully autonomous weapons”.⁴⁹ In fact, the very idea of a machine endowed with the power to take life-or-death decisions is intuitively “repugnant”.⁵⁰ While the basic claim underlying this argument is straightforward (“there is a violation of dignity when a machine kills a human being”), it is informative to distinguish analytically between two of its variants, one of which is centered on agent-relative duties, and the other one on patient-relative rights.

The argument based on human dignity was first put forward by Peter Asaro in a leading article published in 2012 on the *International Review of the Red Cross*.⁵¹ In its original formulation, this argument emphasizes agent-relative duties flowing from the principle of human dignity. The starting point of this argument is that the action of suppressing a human life is legally or morally justifiable only if it is non-arbitrary. However, for an act of taking human lives to be non-arbitrary, the agent should ground his/her action “on a considered and informed decision”.⁵² That is where the principle of human dignity comes in. The latter, in fact, dictates that the act of killing be based on *human* judgment, for only human decision-making guarantees the full appreciation “of the value of individual life [and] the significance of its loss.”⁵³ Therefore, to the extent that lethal decision-making is carried out by a machine without any involvement of human judgment, AWS are inherently unable to comply with this dictate. Accordingly, AWS taking human life are neither legitimate nor morally justifiable.⁵⁴

⁴⁷ Art. 1(2) AP I “[...] civilians and combatants 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” (emphasis added). The issue of compatibility of robotic killings with the Martens Clause was raised at the CCW Meetings of Experts by the following States: Austria (2015 CCW Meeting of Experts, Working Paper on “The concept of ‘meaningful human control’”, p. 2), Brazil (2014 CCW Meeting of Experts, Opening Statement, p. 2), Ecuador (2016 CCW Meeting of Experts, General Exchange, p. 2), and Mexico (2014 CCW Meeting of Experts, Opening Statement, p. 1).

⁴⁸ C. Heyns *cit. supra* note 29, p. 367. See also, in a critical perspective, D. Birnbacher “Are autonomous weapon systems a threat to human dignity?”, in N. Bhuta et al., *cit. supra* note 5, p. 105 ff., pp. 105-106.

⁴⁹ United Nations Institute for Disarmament Research (UNIDIR), *The Weaponization of Increasingly Autonomous Technologies: Considering Ethics and Social Values* (UNIDIR, 2015), p. 7.

⁵⁰ *Ibid.*, p. 8.

⁵¹ P. Asaro, “On banning autonomous weapon systems: human rights, automation, and the dehumanization of lethal decision-making”, *International Review of the Red Cross* (2012), p. 687 ff.

⁵² *Ibid.*, p. 689.

⁵³ Human Rights Watch, *Shaking the Foundations The Human Rights Implications of Killer Robots* (HRW 2014), p. 3.

⁵⁴ It was aptly noted that the prohibition flowing from the principle of human dignity might be independently endorsed from a consequentialist standpoint in normative ethics on account of its expected consequences as a behavioral rule. Indeed, by requiring the involvement of human judgment in life-or-death decision-making may selectively induce a human agent, whose motives comprise compassion and empathy, to refrain from using lethal force, notwithstanding she or he is legally entitled to do so (R. Geiss, *cit. supra* note 24, p. 18). If compassion and empathy do not act as a restraint, “greater killing and suffering in war than would otherwise be necessary” could ensue (A. Saxton, “(Un)Dignified Killer Robots? The Problem with the Human Dignity Argument”, *Lamfare* (20 March 2016), available at <https://www.lawfareblog.com/undignified-killer-robots-problem-human-dignity-argument>). And if AWS “always apply deadly force within what is legally permissible the possibility exists that people will die who otherwise would have been spared” (R. Geiss *cit. supra* note 24, p. 18). At the CCW Meetings of Experts, this point was made, effectively enough, by the delegate of the Holy See, who observed that:

“Studies of soldiers’ experiences support that human beings are innately averse to taking life, and this aversion can show itself in moments of compassion and humanity amidst the horrors of war. Programming an “ethical governor”

The variant of the argument based on human dignity which is centered on patient-relative rights was clearly formulated by the Special Rapporteur on extrajudicial, summary or arbitrary executions, Christof Heyns. His argument is grounded in the Kantian conception of human dignity, which entails – ultimately – that people have the inherent right to be treated as “complete and unique human beings”, especially when their lives are at stake.⁵⁵ Accordingly, human dignity would be blatantly denied if people were subject to robotic lethal decision-making, because this would place them “in a position where an appeal to the humanity of the person on the other side is not possible”.⁵⁶ Indeed, the decision to kill or not would be taken “on the basis of hypotheticals”, which would be set in advance, in the AWS programming phase,⁵⁷ or would be developed by the machine itself as rules of behavior extrapolated from its past experience, provided that the machine had been suitably endowed with some learning algorithm for this purpose. 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 slightest) hope of changing his/her fate. In this way, hope, which constitutes a consubstantial part of human nature, would become “one of the casualties of AWS”.⁵⁸

3.4 The scope of deontological arguments against AWS

Deontological arguments suffer from some inherent limitations in scope, which must be duly taken into account in order to assess their capability to support the case for a comprehensive ban on AWS.

First, while no educated guess can be sensibly advanced as to whether and when AWS will meet IHL, one cannot *ipso facto* exclude the technological possibility of IHL-compliant AWS.⁵⁹ Accordingly, deontological arguments drawing on current and foreseeable violations of IHL by AWS do offer a strong support for a moratorium on the development and use of lethal AWS *in those warfare scenarios where significant problems of distinction and proportionality arise*.⁶⁰ But they are not sufficient to support an

or “artificial intelligence” to enable autonomous weapon systems to technically comply with the law of war in the areas of distinction and proportionality, even if possible, is not sufficient.” (2014 CCW Meeting of Expert, Opening Statement, p. 2).

⁵⁵ C. Heyns, *cit. supra* note 29, p. 370.

⁵⁶ *Ibid.*

⁵⁷ *Ibid.*

⁵⁸ *Ibid.*, p. 371.

⁵⁹ At the CCW Meetings of Experts, this point was highlighted by several EU States: Czech Republic (2014 CCW Meeting of Expert, Statement at the Panel on “Operational and Military Aspects”, p. 2), Finland (2015 CCW Meeting of Expert, Statement at the Panel on “Transparency and the Way Ahead”, pp. 1-2), Greece (2015 CCW Meeting of Expert, Statement at the Panel on “Possible Challenges to IHL”, p. 1), Finland France (2015 CCW Meeting of Expert, Statement at the Panel on “Overarching Issues”, p. 1), Italy (2016 CCW Meeting of Expert, Statement at the Panel on “Towards a Working Definition of Lethal AWS”, p. 2), Spain (2014 CCW Meeting of Expert, Opening Statement, p. 2),

⁶⁰ R.C. Arkin, “Lethal Autonomous Systems and the Plight of the Non-combatant”, *Ethics and Armed Forces* (2014), p. 3 ff., p. 7. See also M.N. Schmitt and J.F. Thurnher, “Out of the Loop?: Autonomous Weapon Systems and the Law of Armed Conflict”, *Harvard National Security Journal* (2013), p. 231 ff., T. Hattan, *cit. supra* note 29, pp. 1048-1051.

unconditional ban on AWS. For all we know, lack of IHL compliance is a contingent situation which may no longer hold at some future time, and the same may be said with regard to IHRL.

Second, the arguments concerning the AWS-related accountability gap do not take into account scenarios where war crimes and crimes against humanity cannot simply be committed, such as aerial and submarine duels between AWS taking place with no humans in their range of action or AWS' attacks directed solely against military communication satellites and other uninhabited military infrastructures. Similarly, the Asaro-Heyns argument addresses problems arising from the *lethal* use of AWS only. Therefore, these arguments offer support for the introduction of a prohibition to employ AWS in those warfare scenarios, which are admittedly crucial for ethical and legal reflections on war, where human lives are at stake. However, they are silent about, say, AWS attacking uninhabited infrastructures or about warfare scenarios involving AWS only.⁶¹

To sum up. A distinction was introduced between deontological arguments which depend on the failure of present and foreseeable AWS to comply with IHL and deontological arguments which do not depend on this contingency, but rather hinge on the need to avoid accountability gaps and to respect the right to human dignity. Deontological arguments of both kinds concern what human agents of certain special sorts that are involved in warfare scenarios morally owe to other human beings that are involved in the same scenarios as potential patients of aggressive actions, in addition to the related question whether these moral and legal obligations would be necessarily or likely to be violated using AWS against human beings. Therefore, these arguments do not bear directly on the problem of prohibiting the use of AWS against uninhabited military infrastructures and in actions against other AWS. Accordingly, there are some limitations of deontological arguments for banning AWS, as they are solely concerned with AWS lethal uses. Consequentialist arguments for an AWS ban, that we now turn to consider, address moral and legal issues concerning AWS, including issues about non-lethal uses of these weapons that are not in the scope of the deontological arguments examined so far.

4. Consequentialist perspectives on banning AWS

Consequentialist approaches in normative ethics assess moral obligations and permissions exclusively on the basis of an evaluation of the (actual or expected) consequences of actions. Accordingly, consequentialist appraisals may conflict with deontological appraisals of AWS deployment. Indeed, depending on the criteria that one adopts to evaluate action consequences, some AWS attacks on humans may be permissible in view of actual or expected consequences that are deemed to be good; and AWS attacks on uninhabited military objectives – which may be morally permissible from the

⁶¹ M.N. Schmitt, *cit. supra* note 9, p. 11 (“[n]ot every battlespace contains civilians”).

deontological viewpoint of agent-relative obligations and patient-relative rights – might be forbidden instead from a consequentialist viewpoint. Thus, the range of moral judgments that one may advance from a consequentialist perspective differs in significant ways from the range of moral judgments one may offer on the basis of the deontological arguments examined in the previous section. In the next sub-sections, we will focus on consequences of AWS deployment with respect to various conceptions of what are good and bad consequences in the realm of armed conflicts and wars. Again, as in the case of deontological arguments, normative ethics and international law approaches intersect. In this case, however, the crucial intersections concern more directly the legal regime governing the maintenance of peace and security rather than IHL and IHRL.

4.1. *Narrow vs. wide consequentialist arguments*

To begin with, let us note that any consequentialist argument for or against a ban on AWS is presently bound to focus on expected, rather than actual, consequences of their deployment since only relatively unsophisticated AWS (see section 2 above) from a much wider spectrum of possibilities are in actual use.⁶² Thus, the consequentialist debate about an AWS ban is more appropriately framed as a *rule-consequentialist* debate concerning which one of two incompatible rules (prohibiting or permitting AWS) must be adopted by the international community of state actors in view of their *expected* outcomes. Accordingly, the debate has essentially developed in terms of different appraisals of the expected consequences of AWS future deployment. In this connection, an informative distinction can be made between wide and narrow appraisals of the expected consequences of AWS deployment.⁶³

Narrow approaches to the appraisal of expected consequences concentrate on expected AWS battlefield performances and some of their outcomes. Wide approaches take into account, in addition to local battlefield implications, expected geopolitical consequences of AWS deployment, which range from regional destabilization and conflicts to global implications on arms races and conflict deterrence.⁶⁴ Thus, wide consequentialist approaches differ from narrow consequentialist approaches in that they evaluate AWS deployment in the light of its expected impact over and above battlefield operations, taking into account lower disincentives to start war, a new arms race, regional and global destabilization threats – going up to and including nuclear deterrence destabilization.

⁶² It is interesting to mention, in this connection, the opinion of the Office of the US Air Force Chief Scientist: “it is possible to develop systems having high levels of autonomy, but it is the lack of suitable V&V [validation and verification] methods that prevents all but relatively low levels of autonomy from being certified for use”. Office of the US Air Force Chief Scientist, *Technology Horizons: A Vision for Air Force Science and Technology 2010–30*, vol. 1, AF/ST-TR-10-01 (Air University Press/Air Force Research Institute: Maxwell AFB, AL, Sep. 2011), p. XX.

⁶³ G. Tamburrini, *cit. supra* note 5, pp. 137-141.

⁶⁴ A wide consequentialist approach is pursued in, e.g., J. Altmann, “Arms control for armed uninhabited vehicles: an ethical issue”, *Ethics and Information Technology* (2013), p. 137 ff..

Those who defend the rule that AWS should be permitted on consequentialist grounds have usually assumed a narrow perspective on expected consequences. Roughly speaking, AWS development and deployment should be permitted insofar as these new conventional arms are expected to bring about reduced casualties in one's own and the opponents' camp, as well as among non-belligerents who happen to be present on the battlefield. This expectation is grounded in the belief that AWS will be capable of performing more accurate targeting than human soldiers, and will be programmed to adopt more conservative decisions to fire insofar as these machines can be made free from human self-preservation concerns.⁶⁵ The force of these narrow consequentialist arguments for the future deployment of AWS depends on another crucial assumption. This is the *ceteris paribus* assumption that the deployment of AWS will not have a significant impact outside battlefield scenarios. However, the weakness of this *ceteris paribus* assumption has been convincingly and repeatedly brought out.⁶⁶ Indeed, one may reasonably expect that the spreading of AWS will bring about comprehensive and long-term consequences for international security along with local and short-term military advantages on the battlefield. A document produced by ICRAC (International Committee for Robot Arms Control)⁶⁷ summarizes various threats to international security raised by AWS concentrating in particular on the proliferation of these weapons with oppressive regimes⁶⁸ and terrorists,⁶⁹ their mass proliferation among state actors giving rise to a new arms race,⁷⁰ less disincentives to start wars,⁷¹ on account of the reduced numbers of soldiers that will be involved, and correspondingly lowered thresholds for armed conflicts, unpredictability of interaction with friendly or enemy AWS, their cyber vulnerability possibly

⁶⁵ See, generally, R.C. Arkin, *Governing Lethal Behavior in Autonomous Robots* (CRC Press, 2009).

⁶⁶ This weakness is aptly emphasized, e.g., in the 2015 "Open Letter on Autonomous Weapons from AI and Robotics Researchers" (<http://futureoflife.org/open-letter-autonomous-weapons/>) and in S. Russell, "Take a Stand on AI Weapons", *Nature* (28 May 2015), pp. 415-416.

⁶⁷ ICRAC, LAWS: Ten Problems for Global Security. Memorandum for delegates at the 2015 Convention on Certain Conventional Weapons (CCW) Meeting of Experts on Lethal Autonomous Weapons Systems (LAWS). <http://icrac.net/wp-content/uploads/2015/04/LAWS-10-Problems-for-Global-Security.pdf>

⁶⁸ HRW, *Losing Humanity*, *cit. supra* note 9, p. 38 ("Fully autonomous weapons would conversely be perfect tools of repression for autocrats seeking to strengthen or retain power. Even the most hardened troops can eventually turn on their leader if ordered to fire on their own people. A leader who resorted to fully autonomous weapons would be free of the fear that armed forces would rebel. Robots would not identify with their victims and would have to follow orders no matter how inhumane they were."); M. Sassòli, "Autonomous Weapons and International Humanitarian Law: Advantages, Open Technical Questions and Legal Issues to be Clarified", *International Law Studies* (2014), p. 308 ff., p. 317 ("if autonomous weapons fall into the hands of a ruthless leader able to program them, they will never abandon his or her cause as soldiers might because the programmed robots will be unable to come to the conclusion that it is unjust, even when there are repeated and widespread violations of IHL.").

⁶⁹ Serious concerns about the possible proliferation of AWS among terrorist groups were expressed at the CCW Meetings of Experts. See the observations made by Algeria (2016 Meeting of Experts, General Exchange, p. 3), Austria (2014 Meeting of Experts, Opening Statement, pp. 1-2), Cuba (2015 Meeting of Experts, General Exchange, p. 1), Ecuador (2016 Meeting of Experts, General Exchange, p. 4), the Holy See (2016 Meeting of Experts, Working Paper "Elements Supporting the Prohibition of Lethal Autonomous Weapons Systems", p. 2), Pakistan (2014 Meeting of Experts, Opening Statement, p. 2), Sierra Leone (2016 Meeting of Experts, General Exchange, p. 2),

⁷⁰ At the CCW Meetings of Experts, this risk has been underscored by Algeria (2016 Meeting of Experts, General Exchange, p. 3), Austria (2015 CCW Meeting of Experts, General Exchange, p. 2), Cuba (2015 Meeting of Experts, Statement at the Panel on "Overarching Issues", p. 2), and the Holy See (2016 CCW Meeting of Experts, Working Paper "Elements Supporting the Prohibition of Lethal Autonomous Weapons Systems", pp. 2-3),

⁷¹ See *infra* section 4.1.

leading to unintended conflicts, acceleration in the pace of battle, in addition to continuous global battlefields brought about by AWS left behind to patrol post-conflict zones over long time periods.

In addition to these concerns, one should carefully note that AWS, more than many other conventional arms, have the potential to deliver destructive attacks on nuclear objectives. Large swarms of AWS flying at supersonic and hypersonic speeds might be capable of delivering a powerful first strike against the opponent's nuclear arsenals, to the extent that they may thwart the opponent's capability of responding with nuclear retaliation. In this scenario, nuclear deterrence based on mutually assured destruction would no longer count as a motivation to withhold aggression and first strike strategies would be prized instead.⁷²

What is then a wide consequentialist appraisal of the overall expected benefits and costs flowing from AWS deployment? Arguably, by permitting the AWS deployment, one might expect the good consequence of reduced casualties among belligerents and non-belligerents in some local battlefield scenarios. By taking this course of action, however, one would significantly raise at the same time the danger of starting a new arms race leading to regional and global destabilization risks, up to and including the weakening traditional nuclear deterrence factors based on mutually assured destruction. As the latter negative consequences outweigh the sum of the expected benefits flowing from AWS deployment, the collective rule of behavior that is expected to produce the preferable set of consequences in a global geopolitical context is that of prohibiting – rather than permitting – the production and deployment of AWS.

4.1 Consequentialist pro-ban arguments in a legal perspective

In the framework of international law debates, consequentialist approaches have been equally pursued by those backing or else opposing a ban on AWS. Significantly enough, the two sides of the debate are positioned coherently with the distinction, set out above, between narrow and wide consequentialist arguments. In fact, the consequentialist approach is commonly wielded by the anti-ban front, which has argued that AWS' deployment will ultimately result in "higher-than-human" performances with respect to adherence to IHL, because robots 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).⁷³

⁷² See G. Tamburrini, *cit. supra* note 5, pp. 139-140.

⁷³ M.N. Schmitt, *cit. supra* note 9, p. 23; K. Anderson and M. Waxman, "Law and Ethics for Autonomous Weapon Systems. Why a Ban Won't Work and How the Laws of War Can" (Hoover Institution, 2013), p. 14; M. Sassòli, *cit. supra* note 68, p. 310

As noted above, however, this narrow appraisal only captures a fraction of the overall picture, since it is confined to the battlefield-related effects and screens off (by the implicitly assumed *ceteris paribus* clause) more pervasive effects that are likely to flow from AWS deployment. Indeed, supporters of a ban reach opposite conclusions on the basis of a broader consideration of the consequences that one may expect from an increased use of AWS. This enlarged perspective brings into play a distinct legal regime, since one moves from the law regulating the conduct of hostilities (IHL, or *jus in bello*) to the law pertaining to the maintenance of international peace and security (Art. 1(1) of the UN Charter). The latter includes, but is not limited to, the rules governing the use of force, or *jus ad bellum* (viz. the prohibition on the use of force under Art. 2(4) of the UN Charter; the right to self-defence under Art. 51 of the UN Charter; and the collective security system governed by Chapter VII of the UN Charter).⁷⁴

A preliminary observation on the shifting legal framework is in order here. While the potential impact of AWS on international peace and security is often described as a matter of concern for international law, this proposition has been rarely discussed in depth, so that is not entirely clear what is the actual legal issue at stake. This led one author to radically 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”.⁷⁵

The latter view is not without foundation. Whether a certain use of force is contrary to the *jus ad bellum* ultimately depends on the circumstances in which force is unleashed (Who? Against whom? Why?) and not on the sorts of weapons that are employed. If, for instance, State A deploys a swarm of AWS against State B, the legality of its conduct will be gauged, under *jus ad bellum*, on the basis of the following elements: whether State A acted in self-defense or whether the use of force was authorized by the UN Security Council. Conversely, it will be completely immaterial whether the attack was carried out through AWS or other alternative means.

Yet, and again, this is not the whole story. The law governing the maintenance of international peace and security cannot be reduced to a static, binary decision 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”.⁷⁶ 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,

⁷⁴ P. Asaro, “How just could a robot war be?”, in A. Briggle, K. Waelbers and P.A.E. Brey (eds), *Current Issues in Computing And Philosophy* (IOS Press 2008), p. 50 ff., pp. 56-58; P. Alston, *cit. supra* note 24, p. 44; HRW, *Losing Humanity*, *cit. supra* note 9, pp. 39-41; C. Grut, *cit. supra* note 21, pp. 22-23; R. Geiss, *cit. supra* note 24, pp. 12-13.

⁷⁵ M. Wagner, *cit. supra* note 33, para. 11.

⁷⁶ A/RES/39/11 (12 November 1984), para. 3.

represent a factor of instability at global and regional levels. Should the latter be the case, any such policy would not only be undesirable as a matter of normative ethics, but also as a matter of international law, as it would run contrary to the maintenance of international peace and security, namely, according to a commonly shared view, it would run counter to the “purpose of all purposes” of the UN Charter.⁷⁷

In the previous sub-section, a variety of nefarious wide-scale consequences were listed, which are likely to ensue from permissive policy towards AWS. Each one of these consequences, taken individually, is arguably sufficient to support the contention that an AWS permissive policy should be outlawed as detrimental to the achievement of the UN goal of a world order of peace and security. Here, we limit ourselves to underlining that a policy allowing the use of AWS would end up encouraging a more liberal approach to the use of force by States. In their turn, such liberal approaches may bring about a higher likelihood of violations of the prohibition on the use of force under Art. 2(4) of the UN Charter.

The present argument is based on a straightforward assumption. Even if one does not fully embrace the (admittedly controversial) democratic peace theory, one must acknowledge that, in democratic countries, public opinion and legislative assemblies play an important role in deterring governments from deploying their armed forces in aggressive military campaigns. In this respect, a crucial factor lies in the risk of casualties among national military personnel. Indeed, popular outrage generally stemming from the return of “flag-draped coffins” represents a significant incentive for representatives sitting in parliaments to exert a meaningful control over the use of war powers by the executive. As a collateral (and not necessarily intended) effect, this democratic dynamics may prevent States from breaching the prohibition on the use of force. A notable case in point occurred in 2013, when the US and UK governments renounced to wage (an arguably unlawful) war against Assad in Syria apparently in view of the disapproval expressed by domestic public opinion and parliamentary representatives.⁷⁸

A policy allowing for the use of AWS would inevitably affect this virtuous circle. If human troops are replaced, say, by robots, the potential cost of the conflict in terms of human losses significantly decreases (when it does not equate to zero) and, with it, sensitivity to the issue in the general public. Accordingly, legislative assemblies would be less motivated to control the governmental exercise of war powers, thereby encouraging further executive unilateralism in this field.⁷⁹ As a final result,

⁷⁷ R. Wolfrum, “Ch. I Purposes and Principles, Article 1”, in B. Simma et al. (eds.), *The Charter of the United Nations: A Commentary* (OUP, 3rd ed, 2012), Vol. I, p. 108 ff., p. 109.

⁷⁸ H. Dieck, “The United States and the Syrian crisis: The Influence of Public Opinion on the Non-Intervention Policy” (29 January 2014), available at <http://www.sciencespo.fr/cei/fr/content/united-states-and-syrian-crisis-influence-public-opinion-non-intervention-policy>; A. Grice “Syria crisis: The British public has its say as two-thirds oppose strikes”, *The Independent* (3 September 2013), available at <http://www.independent.co.uk/news/uk/politics/syria-crisis-the-british-public-has-its-say-as-two-thirds-oppose-strikes-8795319.html>.

⁷⁹ For discussion of this issue, especially in connection with the use of drones in the US 2011 Libya intervention, see E. Datteri and G. Tamburrini, “Robotic Weapons and Democratic Decision-Making”, in E. Hilgendorf and Günther J-P (eds.), *Robotik und Gesetzgebung* (Nomos Verlagsgesellschaft 2013), p. 211 ff.

democratically unchecked military operations will be more and more likely to occur, leading to more frequent breaches of the prohibition on the use of force. As the Austrian delegation openly put it 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.”⁸⁰

Significantly enough, a strikingly similar conclusion has been reached by Rebecca Crootof, a resolute critic of an AWS ban. In a recent analysis devoted to the influence of AWS on the distribution of war powers in the US legal system, Crootof describes how (and why) the development of this technology would lead to a further concentration of the war power in the Executive’s hands.⁸¹ When turning to consider the international legal implications of this process, she notes that, as a result of this concentration of power, US Presidents “will be more willing to engage in humanitarian interventions”.⁸² At this juncture, the author acknowledges that unilateral humanitarian interventions are still prohibited under international law, but suggests that a more frequent resort to military force by the US could lead to the consolidation of a new exception to the prohibition on the use of force under Art. 2(4) of the UN Charter.⁸³ Whether this result could ever be achieved, however, it is highly doubtful, if one only considers that substantial portions of the world community have treated US-led unilateral humanitarian interventions with the utmost suspicion. What is less controversial, for the time being, is that more frequent resort to military force will ultimately mean more frequent violations of the prohibition on the use of force. Accordingly, Crootof’s statement quoted above is aptly turned on its head and rephrased as follows: because of the concentration of power stemming from the use of AWS, US Presidents “will be more willing to engage in *violations of the prohibition on the use of force under Art. 2(4) of the UN Charter*”. This amended conclusion affords a quite strong motivation for banning AWS!

4.3. The scope of consequentialist arguments against AWS

If taken in isolation, consequentialist arguments – similarly to deontological ones – suffer from some inherent limitations, which must be duly taken into account in order to assess their force in support of a comprehensive ban on AWS.

⁸⁰ Opening Statement, pp. 1-2. See also the views expressed by Costa Rica (2016 CCW Meeting of Experts, General Exchange), Chile (2015 CCW Meeting of Experts, General Exchange, p. 2), Cuba (2015 CCW Meeting of Experts, General Exchange, p. 1), Ecuador (2016 CCW Meeting of Experts, General Exchange, p. 2), France (2015 CCW Meeting of Experts, Statement at the Panel on “Overarching Issues”, p. 2), the Holy See (2015 CCW Meeting of Experts, Statement at the Panel on “Overarching Issues”, p. 3), Pakistan (2014 CCW Meeting of Experts, Opening Statement, p. 2), Sri Lanka (2015 CCW Meeting of Experts, General Exchange, p. 2,)

⁸¹ R. Crootof, “War, Responsibility, and Killer Robots”, *North Carolina Journal of International Law and Commercial Regulation* (2015), p. 909 ff.

⁸² *Ibid.*, p. 929.

⁸³ *Ibid.*, pp. 930-931.

First, AWS proliferation with oppressive regimes and terrorist groups could be effectively countered by the introduction of an *ad hoc* control legal regime (either in the form of a treaty or of a Security Council resolution), aimed at strengthening international cooperation to prevent AWS from ending up in the “wrong” hands.⁸⁴

Second, not every use of AWS by States is likely to yield destabilizing effects. Just consider, for instance, AWS performing merely defensive (e.g. the Israeli Iron Dome) or sentinel functions (such as the gun turret installed along the demilitarized zone between South and North Korea or, more recently, along Turkish borders). In this perspective, one might argue, a comprehensive ban would appear as an excessive legal response. Indeed, States and NGOs might alternatively negotiate a prohibition on the production and use of AWS for aggressive and potentially destabilizing purposes, leaving aside, e. g., those carrying out defensive and sentinel functions, which represent a substantive portion of AWS that are currently in use.

In summary, by sufficiently enlarging the temporal and spatial horizon of what must be included in the list of expected consequences, it appears that a policy allowing, in general terms, the use of AWS is likely to produce negative destabilizing effects that largely outweigh positive ones. However, this argument, if taken in isolation, provides limited support only for banning AWS, since the expected destabilizing effects deriving from an increasing spread and use of AWS might be successfully offset through the adoption of various restrictive measures falling short of a sweeping ban.

5. Deontological and consequentialist reasons for banning AWS: a confluence model

Let us recap the main points of the analysis developed in the previous sections.

- Deontological arguments raise strong objections to *AWS targeting humans*, but afford weaker and only indirect arguments for AWS not targeting humans.⁸⁵
- Wide consequentialist arguments raise strong objections against AWS, and also against those AWS that are not directed towards human targets, *as long as there are expected destabilizing effects at various scales*. However, if no destabilizing effects are expected, and good consequences are in the offing without being counterbalanced by bad consequences, consequentialist arguments may lead one to support the introduction of AWS for targeting humans in view of their greater targeting accuracy and more conservative firing decisions.

⁸⁴ Sassòli, *cit. supra* note 68, p. 317.

⁸⁵ In a “slippery slope” perspective, for instance, it may be argued that, if the use of AWS non targeting humans is allowed, the possibility remains that a commander could feel compelled, out of military necessity, to employ such AWS in a scenario involving the targeting of human beings.

Accordingly, there is a tension between deontological and consequentialist approaches to the issue of AWS deployment and use, which leads to conflicting conclusions about lethal uses of AWS having no evident destabilizing effects. There is, however, elbow room for reconciling these different approaches and to bolster on this basis the case for an AWS ban. To begin with, let us notice that deontological arguments in normative ethics concern some agent-relative obligations and patient-relative rights in warfare scenarios involving the use of lethal force by means of AWS. Moreover, the conclusions of deontological arguments against AWS flow from inviolable foundational values (human dignity) and requirements of moral responsibility and accountability that are deeply entrenched into IHL, IHRL and ICL. Under these bodies of law, rules concerning the use of lethal force are categorical and cannot be derogated from in the pursuit of a greater good (as a consequentialist analysis would suggest), except for special excusing conditions that are explicitly envisaged by the rules themselves.⁸⁶ The joint ethical and legal reinforcement of categorical obligations motivates the following Prioritization of Deontology rule (PD-rule): the conclusions of deontological arguments cannot be overridden in any circumstance in which they are applicable. However, it was noticed that there are circumstances in which deontological arguments do not provide such guidance, i.e. when the obligations of certain sorts of agents or the rights of certain sorts of potential patients are not at stake. In these circumstances, deontological approaches – both ethical and legal ones – are inapplicable. How is this deontologically unregimented space of action possibilities to be dealt with? Here, the agent-neutral (and patient-neutral) consequentialist framework can provide the required guidance, insofar as wide consequentialist reasons agree with and are enshrined into the principles (set forth by the UN Charter and related legal instruments) concerning peace and security. In the deontologically unregimented space of action, wide consequentialist reasons are reinforced by a variety of legal instruments concerning peace and security. Thus, deontological and consequentialist frameworks can be amalgamated by assigning them to different domains, thereby avoiding intertheoretical conflicts in normative ethics and legal conflicts between different bodies of law.

The normative model obtained by merging different ethical and legal frameworks under the PD rule (which prioritizes deontological reasons whenever applicable and uses consequentialist reasons otherwise), supports the prohibition of using AWS against human beings *and* in any other circumstance which may give rise to threats for peace and destabilization effects at various scales. This model (PD-confluence model from now on) compensates the respective strengths and limitations of deontological and consequentialist reasons that have been offered in public debates and scholarly contributions on the prohibition of AWS deployment and use. For this reason, the PD-confluence model bolsters the case for an AWS ban more effectively than deontological and consequentialist reasons taken in

⁸⁶ See, generally, W.P. Gormley, “The right to life and the rule of non-derogability: peremptory norms of *jus cogens*” in B.G. Ramcharan (ed.), *The Right to Life in International Law* (Martinus Nijhoff, 1985), p. 120 ff.

isolation from each other; but also more effectively than a superficial juxtaposition of deontological and consequentialist reasons which disregards intertheoretical tensions in normative ethics and consistency issues between different bodies of law.

In conclusion, let us briefly outline some implications of the PD-confluence model for the problem of assessing what is the meaningful human control (MHC) that humans have to exercise on weapons systems. This notion, introduced in the public debate on AWS by the ONG Article 36⁸⁷, has been used by Human Rights Watch to define AWS as those “weapons systems that would select and engage targets *without meaningful human control*”.⁸⁸ While there is a growing consensus on the need to subject all weaponry to MHC,⁸⁹ it is highly controversial what MHC is actually supposed to mean.⁹⁰

Our working hypothesis is that the PD-confluence model provides guidance as to the MHC one has to apply for a weapon not to be regarded as “autonomous” under a future AWS treaty. To illustrate these implications of the PD-confluence model, let us preliminarily recall the following basic construals of the notion of MHC (which, of course, may be further refined and specified):

1. MHC requires appropriate system programming and testing of an AWS⁹¹
2. MHC requires the ability of a human operator to oversee and veto an AWS⁹²
3. MHC requires deliberative moral reasoning, by human beings, over individual attacks of an AWS⁹³

⁸⁷ See the Article 36 briefing paper “Key Areas for Debate on Autonomous Weapons Systems” (May 2014), <http://www.article36.org/wp-content/uploads/2014/05/A36-CCW-May-2014.pdf>. For recent technical discussion on MHC, see M.C. Canellas and R.A. Haga, “Toward Meaningful Human Control of Autonomous Weapons Systems through Function Allocation”, *IEEE International Symposium on Technology in Society (ISTAS) Proceedings* (2015),

⁸⁸ HRW, *Mind the Gap*, *cit. supra* note 39, p. 6.

⁸⁹ This led one author to argue that MHC would be imposed by general international law. See P. Asaro, “*Jus nascendi*, robotic weapons and the Martens Clause”, in R. Calo, A.M. Froomkin and I. Kerr (eds.), *Robot Law* (Elgar, 2016), p. 367 ff. Likewise, some States contended that MHC over weaponry would be required under IHL. This view was expressed by Austria (2014 CCW Meeting of Expert, Closing Statement), Croatia (2014 CCW Meeting of Expert, Closing Statement), Germany (2014 CCW Meeting of Expert, Opening Statement, p. 4), and Sri Lanka (2015 CCW Meeting of Expert, General Exchange, pp. 2-3).

⁹⁰ R. Crotoft, “The Meaning of ‘Meaningful Human Control’”, *Temple International & Comparative Law Journal* (forthcoming), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2705560, p. 3. In a similar vein, the UK delegation at the 2016 CCW Meeting of Experts noted that “what may or may not be meaningful is almost an entirely subjective judgment” (Statement at the Panel on “Towards a Working Definition of Lethal AWS”, p. 2). Strong criticism was also raised by the French delegation, according to which the concept of MHC “contradicts the very notion of full autonomy and lacks the precision and technical accuracy needed to discuss complex prospective systems” (2016 CCW Meeting of Experts, Non Paper on “Characterization of a Lethal AWS”, p. 2). See also the views expressed by Finland (2015 CCW Meeting of Expert, Statement at the Panel on “Transparency and the Way Ahead”, p. 1), India (2015 CCW Meeting of Expert, Statement at the Panel on “Transparency and the Way Ahead”, p. 2), and United States (2016 CCW Meeting of Expert, General Exchange, p. 2).

⁹¹ This definition of MHC, for instance, was endorsed by the Israeli delegation at the 2015 and 2016 CCW Meetings of Experts (see, respectively, Statement at the Panel on “Characteristics of Lethal AWS”, p. 2; and General Exchange, p. 3).

⁹² See, in this connection, the Statement by the Polish delegate at the Panel on “Human Rights and Ethical Issues” of the 2016 Meeting of Experts (“[...] we would like to propose to look at the possibility of human control over the robotic systems rather than the actual execution of such control. Following this logic, a person accountable for robot actions is the user who has a possibility to take over control over a robotic system at every moment of the robot conduct, without necessarily executing such control”).

We contend that the issue of MHC does not lend itself to a one-size-fits-for-all solution which can be achieved, e.g., by endorsing one and only one of (1)-(3).⁹⁴ Nevertheless, we maintain that the “appropriate levels of human control”⁹⁵ in each case must be determined on a principled ethical and legal basis. One coherent principled basis is afforded by the deontological and consequential arguments that are amalgamated and prioritized in accordance with the PD-C model. To illustrate, let us consider some of the existing weapons systems mentioned in section 2 above, which can be classified as autonomous on the basis of the DoD requirements.

- In the case of Israeli Iron Dome (and similar anti-materiel defensive systems), the time-frame operational constraints do not allow for real-time judgments by officers. However, MHC at the level of design, implementation and testing of the weapons system appears to be sufficient in the light of the PD model. Indeed, the system is neither directed against humans, nor has destabilizing effects. Moreover, it operates in preset areas to be defended according to informed human judgment by military officers. And if the AWS hits human targets, programmers and manufacturers can be held responsible for their culpable actions or omissions.
- In the case of sentinel gun towers, such as the South Korean SGR-A1 and the Turkish border sentinel, MHC should be understood, under the PD model, as implying human supervision and veto. Unlike the former case of the Iron Dome, operational time-frames are compatible with those human deliberation processes that are needed for a veto to be issued. Moreover, the system is designed to use lethal force against humans, and cannot recognize unconventional surrender gestures. Finally, renouncing to the unsupervised operation mode poses no threats to friendly forces. By not vetoing, the human supervisor will ultimately decide the use of lethal force and will answer for it accordingly.
- In the case of the UK Brimstone (and similar fire-and-forget munitions)⁹⁶, the unleashing of this system in its autonomous mode must be subject to the informed judgment of an operator about each single activation: the operational area and the time-frame of action must be carefully recognized as ones where, according to the operator’s pondered assessment, the fire-and-forget munitions system will not violate distinction and proportionality requirements. If there is any

⁹³ Arguably, this is the definition ensuing from the requirements set forth by Article 36 (*Killer Robots: UK Government Policy on Fully Autonomous Weapons* (April 2013), p. 4) and the Center for a New American Security (M.C. Horowitz and P. Scharre, *Meaningful Human Control in Weapons Systems: A Primer* (CNAS, March 2015), p. 13).

⁹⁴ At the CCW Meetings of Expert, the need for a flexible approach to MHC was advocated by Sweden (2015 CCW Meeting of Expert, General Exchange, pp. 1-2).

⁹⁵ This expression was coined by the US delegation as an alternative to a rigid notion of MHC (2016 CCW Meeting of Expert, General Exchange, p. 2).

⁹⁶ For instance, the anti-radiation drones Harpy and Harpy 2 (or Harop) developed by the Israel Aerospace Industries. For more information, see <http://www.iai.co.il/2013/36694-16153-en/IAL.aspx>.

doubt that the environment is going to change rapidly in ways that raise IHL (or IHRL) issues, then MHC must be guaranteed by adopting the semi-autonomous, laser-guided mode.

Agent-relative obligations and patient-relative rights, which constitute the prioritized backbone of the PD-confluence model, are preserved in these various cases by singling out human beings who are put in charge of exerting the appropriate level of MHC which involves formulating competent judgments and taking informed decisions as to the use of force. Accordingly, by relying on the PD-confluence model one is able, at least in a variety of significant and concretely available AWS, to pursue a principled approach to the problem of imposing appropriate levels of MHC without having to commit oneself *a priori* to one fixed and decontextualized interpretation of what MHC amounts to.