



TEL AVIV UNIVERSITY
THE BUCHMANN FACULTY OF LAW
ANNY AND PAUL YANOWICZ CHAIR
IN HUMAN RIGHTS

אוניברסיטת תל-אביב
הפקולטה למשפטים ע"ש בוכמן
הקתדרה לזכויות האדם
על שם אני ופול ינוביץ'



PROFESSOR EYAL BENVENISTI

פרופסור איל בנבנישתי

GLOBALTRUST RESEARCH PROJECT

GLOBALTRUST WORKING PAPER SERIES 10/2014



THE OBLIGATION TO EXERCISE DISCRETION IN WARFARE: WHY AUTONOMOUS WEAPON SYSTEMS ARE UNLAWFUL

ELIAV LIEBLICH AND EYAL BENVENISTI

THE GLOBALTRUST WORKING PAPER SERIES CAN BE FOUND AT
[HTTP://GLOBALTRUST.TAU.AC.IL/PUBLICATIONS](http://GLOBALTRUST.TAU.AC.IL/PUBLICATIONS)

ISSN 2309-7248

© ELIAV LIEBLICH AND EYAL BENVENISTI

CITE AS: ELIAV LIEBLICH AND EYAL BENVENISTI, THE OBLIGATION TO EXERCISE
DISCRETION IN WARFARE: WHY AUTONOMOUS WEAPON SYSTEMS ARE
UNLAWFUL, GLOBALTRUST WORKING PAPER 10/2014
[HTTP://GLOBALTRUST.TAU.AC.IL/PUBLICATIONS](http://GLOBALTRUST.TAU.AC.IL/PUBLICATIONS)

The Obligation to Exercise Discretion in Warfare: Why Autonomous Weapon Systems are Unlawful

Eliav Lieblich* & Eyal Benvenisti **

Abstract

This Chapter analyzes the current discourse on the legality of autonomous weapons ('killer robots') under international law, and seeks to offer a novel prism through which to discuss the challenges that such systems pose: namely, we advance the view that modern warfare is an exercise of executive power by states against individuals and should thus be subject to basic notions of administrative law, chiefly the obligation to exercise proper administrative discretion. When autonomous weapons are deployed, state power is exercised against individuals through a computerized proxy. When the power to make 'decisions' that affect basic rights is transferred to computer systems such as autonomous weapons, the duty to exercise proper administrative discretion is compromised. Chiefly, this is because at least in the foreseeable future, such machines would be incapable of exercising 'true' administrative discretion. This problem is especially acute when autonomous weapons are deployed in asymmetric settings, where civilians are put at risk, but also, arguably, when deployed against enemy combatants.

INTRODUCTION

I. TECHNICAL VERSUS SUBSTANTIVE AUTONOMY

II. THE CIRCULAR DEBATE ON AUTONOMOUS WEAPONS

A. *The Instrumentalist Discourse*

B. *The Deontological Discourse*

III. REFRAMING THE DEBATE: WARFARE AS AN EXERCISE OF EXECUTIVE POWER

A. *War as Governance: Warfare as an Exercise of Administrative Power*

B. *Autonomous Weapons and the Problem of Pre-bound Discretion*

C. *Subjecting Warfare to Administrative Law Standards and the Problem of Pre-Bound Discretion: Possible Challenges*

D. *Diagonal Obligations towards Combatants, Pre-Bound Discretion and the Shrinking Window of Surrender*

CONCLUSION

INTRODUCTION

The question of ‘killer robots,’ or autonomous weapon systems (AWS), has garnered much attention in recent discourse. While officials often downplay the prospect of such systems making targeting and other crucial decisions, their own statements reveal the possibility that such capabilities would be developed in the near future.¹ For instance, in late 2012, the US Department of Defense (DoD) imposed a *de facto* moratorium on the development of fully AWS, by emphasizing that weapons ‘shall be designed to allow commanders and operators to exercise appropriate levels of human judgment over the use of force.’² From the mere fact that the DoD felt the need to constrain the development and use of AWS for the time being, in several ways,³ we can learn that the prospect of such weapons is realistic. Indeed, the DoD Directive itself includes a bypass clause in which deviations from its requirements can be approved by high-ranking officials through special procedures.⁴ There is therefore a consensus, among commentators, that the motivation to develop and deploy AWS will overcome these temporary constraints.⁵ This makes the discussion of AWS a timely one.

When discussing the legality and legitimacy of such weapons, the claim that machines should not be making ‘decisions’ to use lethal force during armed conflict is an intuition shared by many. However, the current discussion as to just *why* this is so is rather unsatisfying. The ongoing discourse on AWS is comprised

* Assistant Professor, Radzyner Law School, IDC Herzliya.

** Anny and Paul Yanowicz Professor of Human Rights, Tel Aviv University Faculty of Law. The authors wish to thank Itamar Mann, Adam Shinar & Keren Yalin-Mor for their thoughtful insights. We acknowledge the support of the ERC Advanced Grant (grant No 323323).

¹ Human Rights Watch and International Human Rights Clinic at Harvard Law School, ‘Losing Humanity: The Case Against Killer Robots’, (2012).

² Department of Defense Directive No. 3000.09, §4(a) (Nov. 21, 2012) [hereinafter DoD Directive].

³ Beyond insisting that weapons allow for appropriate human judgment, the Directive requires, for instance, that systems do not operate too fast for human scrutiny (§4.a.(1)(b)); that they have understandable user interfaces (§4.a.(3)(a)); importantly, the Directive distinguishes between ‘semi-autonomous’ and ‘human-supervised autonomous’ weapons for the sake of targeting: while the former can be used as lethal weapons, the latter cannot be used for selection and engagement of human targets, and can only be used as static defense of manned installations or platforms, ‘for local defense to intercept attempted time-critical saturation attacks’ (§4.c.)

⁴ *Ibid.* §4.c.1, Enclosure 3.

⁵ See N. Sharkey, ‘The Evitability of Autonomous Robot Warfare’, *International Review of the Red Cross*, 94 (2012) 787–799, 788 (2012); M.N. Schmitt and J.S. Thurnher, ‘“Out of the Loop”: Autonomous Weapons Systems and the Law of Armed Conflict’, *Harvard National Security Journal*, 4 (2013), 232–281, 237.

of nuanced approaches found between two extremes, arguing against each other roughly along consequentialist and deontological lines. On one side of the spectrum is the approach that if AWS could deliver good results, in terms of the interests protected by international humanitarian law (IHL), there is no reason to ban them. On the contrary: as the argument goes, if we take humanitarian considerations seriously, we should *encourage* the development and use of such weapons. Of course, proponents of this approach envision technological advancements that would actually make such results possible. Found on the other side of the spectrum, are those that claim that even if AWS could, in terms of outcomes, adhere to the basic norms of IHL, their use should still be prohibited, whether on ethical or legal grounds. Usually, those holding this position are also skeptical that technology would *ever* be able to produce such benevolent systems. The discussion, thus, is caught in a loop of utilitarian arguments and deontological retorts. In essence, the current debate on AWS is a manifestation of the circular argument between technological optimists and pessimists found in other contexts.⁶

In this Chapter, we do not attempt to prove any approach within this spectrum ‘wrong.’ We likewise do not aim to propose all-encompassing solutions to the problem of AWS. Rather, we aim to suggest a framework that explains better some of the objections to such weapons. We suggest, therefore, another prism through which the question should be analyzed: one flowing from an *administrative* perception of warfare. In essence, we offer to approach the emerging phenomenon of AWS as an exercise of state power against individuals through a computerized proxy: as such, we suggest analyzing such interactions according to basic notions of administrative law, chiefly the obligation to exercise proper administrative discretion.

This understanding bridges between the administrative (or regulatory) realm, on the one hand, and the realm of warfare, on the other. Although, as we will point out, modern IHL already includes an administrative law aspect in its rules concerning target selection and the need to take ‘constant care’ to spare the

⁶ See in another context, J.E. Krier and C.P. Gillette, ‘The Un-Easy Case for Technological Optimism’, *Michigan Law Review*, 84 (1985), 405–429; A.D. Basiago, ‘The Limits of Technological Optimism’, *Environmentalist*, 14 (1994), 17–22; C.E. Harris et al, *Engineering Ethics: Concepts and Cases*, 4th edn. (Wadsworth, 2013) pp. 80–86 (explaining the notions of technological determinism, optimism and pessimism).

civilian population, that aspect has never been explicitly articulated – and hence the two realms were traditionally viewed as unconnected. As we see it, the administrative law view becomes apparent when discussing the legality of AWS, as it provides the missing link between some objections to AWS, and the question of how sovereigns should act during armed conflict. Specifically, the administrative perception can explain why notions of ‘due process’ or ‘procedural justice’ – which are usually reserved to peacetime relations between a state and its citizenry – can be relevant also to the relations between belligerents (and between them and civilians) during armed conflict, and thus affect the legality of the use of AWS. As we claim, administrative law thinking is not simply another way to argue for the application of international human rights law (IHRL) during armed conflict, although these realms are mutually reinforcing. Rather, administrative law adds to our understanding of human rights, by its emphasis on the proper *process* of decision rather than simply on the *outcome* of the decision making.

Discussing the issue of AWS through the administrative prism can thus assist us to understand better and articulate existing objections to AWS, and can serve as an interpretive or complementing source of principles when analyzing positive IHL and IHRL. Importantly, this approach allows us to assess the issue of AWS even under the assumption that such weapons could, in due time, perform reasonably, in terms of immediate *results*.

The discussion proceeds as follows. Part I defines autonomy for the sake of our discussion, laying down the distinction between ‘technical’ and ‘substantive’ autonomy. Part II outlines the ‘circular’ debate on AWS, caught between the instrumentalist and deontological discourses. Part III suggests analyzing the problem of AWS through the prism of administrative law, which exposes key problems relating to their use – both in relation to civilians and, perhaps – and here we make a preliminary argument – also to enemy combatants. Namely, we argue that owing to the administrative-legal duty to exercise constant discretion, the final decision to use lethal force must be taken by a human being in real time; and that furthermore, AWS may present a significant problem by diminishing the option of surrender.

I. TECHNICAL VERSUS SUBSTANTIVE AUTONOMY

Before we delve into our analysis, it is needed to define what we mean when referring to *autonomous* weapons. Indeed, as evident from the recent meeting of the CCW High Contracting Parties on Lethal Autonomous Weapons Systems, any attempt to offer an elaborate definition is slippery, and in any case likely to be highly technical.⁷ Since autonomy is a spectrum rather than an absolute term,⁸ discussing autonomy without offering a simplified definition is like aiming at a moving target. For the purpose of this work, thus, we offer a simple working distinction between *substantive* and *technical* autonomy, and focus on our analysis strictly on the latter.

Substantive autonomy assumes machines with complete ‘sentient’ abilities. If the prospect of such machines was our point of departure, our discussion would quickly cross the lines to a philosophical debate concerning the essential components of humanity. Are sentient machines robots or quasi-humans? If they are, are their grounds to treat them differently in terms of law? We leave such questions to others.⁹ Likewise, our analysis is not concerned with ‘Singularity’ scenarios,¹⁰ in which humanity and technology become enmeshed, nor with apocalyptic ‘Terminator’-type debacles where machines become the dominant ‘life form’ on Earth.¹¹ In this Chapter we are concerned with reasonably foreseeable developments, which assume an interim perception of *technical* autonomy.¹²

⁷ See, e.g., ‘Report of the 2014 Informal Meeting of Experts on Lethal Autonomous Weapons, Advanced Version’, (16 May, 2014), ¶¶19–20; see also A. Krishnan, *Killer Robots: Legality and Ethicality of Autonomous Weapons*, (Ashgate, 2009), pp. 33–61.

⁸ See, e.g., *ibid.* p. 4 (‘[t]he smaller the need for human supervision and intervention, the greater the autonomy of the machine.’).

⁹ See, e.g., N. Haslam et al, ‘Subhuman, Inhuman, and Superhuman: Contrasting Humans with Nonhumans in Three Cultures’, *Social Cognition*, 26 (2008) 248–258.

¹⁰ See generally Ray Kurzweil, *The Singularity is Near: When Humans Transcend Biology*, (Penguin, 2005).

¹¹ See generally Kevin Warwick, *March of the Machines: the Breakthrough in Artificial Intelligence* (University of Illinois Press, 1997).

¹² We are aware that by out-ruling substantive autonomy we too are making a ‘skeptical’ assumption regarding technological developments. However, we do not derive normative conclusions from this skepticism: we simply refer to it as a given, ‘realistic’ scenario.

Technically autonomous weapons, according to a widely accepted definition, are systems that ‘once activated, can select and engage targets without further intervention by a human operator.’¹³ The system therefore possesses the capability for autonomous choice, in the sense that it is able to take ‘targeting decisions’ with humans ‘out of the loop;’¹⁴ however, the definition stops short of assuming the sentience that constitutes *substantive* autonomy.¹⁵ The latter is negated because the human remains in the ‘wider loop’ of the weapons’ actions: humans predetermine its actions – even if not *per se* controlling them – by programming its algorithms.¹⁶

Since a technically autonomous machine is incapable of altering its algorithms through a process that could be equated to human learning, we cannot claim that it engages in a true process of discretion: its ‘inner deliberation’ is controlled by determinations made *ex ante* by humans. Risking oversimplification – and ongoing research notwithstanding – it is safe to say that computers are generally unable to engage in ‘thinking about thinking.’ or *metacognition*.¹⁷ In general, metacognition comprises a ‘knowledge of cognition,’ which refers to

¹³ DoD Directive, p. 13; Human Rights Watch, ‘Losing Humanity’, p. 2; ‘Report of the Special Rapporteur on Extrajudicial, Summary or Arbitrary Executions, Christof Heyns’, UN Doc.A/HRC/23/47 (9 April 2013) ¶38 [hereinafter ‘AWS Report’].

¹⁴ AWS Report, ¶27, 39.

¹⁵ It should be noted that technically autonomous weapons differ from ‘automated’ or ‘semi-autonomous’ weapons, since the latter still operate under direct control, or at least constant real-time supervision. We set such systems aside since if ‘autonomy’ is limited by the constant oversight of a ‘human in the loop,’ then the weapons do not, at large, raise questions different from those raised by remote controlled weapons such as drones. See P. Asaro, ‘On Banning Autonomous Weapon Systems: Human Rights, Automation, and the Dehumanization of Lethal Decision-Making’, *International Review of the Red Cross*, 94 (2012) 687–709, 690; DoD Directive, Glossary; Schmitt and Thurnher, ‘Out of the Loop’, 234–236. It should be born in mind, however, that a valid question is whether automated or semi-autonomous systems create bias in the sense that the human operator believes these are error-proof. See Danielle Keats Citron, ‘Technological Due Process’, *Washington University Law Review* 85, (2008), 1249–1313, 1254, 1271 (discussing literature that highlights that human supervision over machines is ineffective since they tend to believe that machines are error-proof).

¹⁶ AWS Report, ¶39. Autonomous weapons with a human in the ‘wider loop’ may or may not be subject also to a human ‘on the loop’: meaning, a human that can override the robot’s ‘decisions.’ This option was criticized as unrealistic, in the sense that robots make decisions faster than humans, and therefore such ‘supervised autonomy’ would be ineffective. See Autonomous Weapons Report, ¶41. Human Rights Watch, for instance, views systems with humans ‘on the loop’ as fully autonomous. Human Rights Watch, at 2. For this reason, for instance, the DoD requires that any weapons’ system will ‘complete engagement in a timeframe consistent with the commander and operator intentions’ – in order to allow human intervention. DoD Directive, §1(b).

¹⁷ For a general overview see M.T. Cox, ‘Metacognition in Computation: A Selected Research Review’, *Artificial Intelligence*, 169 (2005), 104–141.

several aspects of knowledge of one's *own* cognition or about cognition at large;¹⁸ and 'regulation of cognition,' which refers to how humans control their learning.¹⁹ For our purposes, we presume that both levels are required in order to exercise human-like discretion, *especially* in complex circumstances: a capacity that technically autonomous weapons do not possess.

Before moving on, a clarification must be made. One can argue that by focusing on *human* discretion, we are placing it on an essentialist pedestal: after all, the application, by a computer, of highly developed algorithms could amount, for all practical matters, to an exercise of discretion.²⁰ However, since we focus on 'technical' autonomy, we exclude scenarios in which AWS actually achieve such human-like capacities, and thus refrain from judging them. Nonetheless, the critique from essentialism can be augmented by a determinist outlook – arguing that human discretion too is pre-bound, since there is no such thing as human free action.²¹ Regarding this claim, we simply presume that imputation of responsibility to human beings is possible²² –if only because its negation results in ruling out the mere possibility of normative judgment, and thus of law, to begin with.²³

Thus, our focus on technical autonomy, combined with the realization that such autonomy excludes the exercise of discretion, has both a methodological and normative importance. Methodologically, it allows us to shift the discussion from the metaphysics of substantive autonomy to the more earthly issue of the process of decision-making. Normatively, as we shall see, it has ramifications when considering the use of AWS under notions of administrative law: this is because

¹⁸ G. Schraw, 'Promoting General Metacognitive Awareness,' *Instructional Science*, 26 (1998) 113–125, 114 (1998) (describing 'cognition of knowledge' as consisting of 'declarative,' 'procedural' and 'conditional' knowledge).

¹⁹ *Ibid.*

²⁰ T.J. Barth and E.F. Arnold, 'Artificial Intelligence and Administrative Discretion: Implications for Public Administration,' *American Review of Public Administration*, 29 (1999), 334–336.

²¹ See generally C. Hofer, 'Causal Determinism', in E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (2010), <http://plato.stanford.edu/entries/determinism-causal/>.

²² For a detailed discussion of determinism and moral responsibility see A. Eshleman, 'Moral Responsibility', in E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (2014), <http://plato.stanford.edu/cgi-bin/encyclopedia/archinfo.cgi?entry=moral-responsibility>.

²³ For a classic statement of the constitutive role of imputability in the legal system see Hans Kelsen, *Introduction to the Problems of Legal Theory*, new edn, B. Litschewski Paulson and S.L. Paulson (trans.), (Clarendon Press, 1997), §§15–16.

the rigid structuring of decision-making, through programming, is comparable to the ‘pre-binding’ of administrative discretion.

II. THE CIRCULAR DEBATE ON AUTONOMOUS WEAPON SYSTEMS

As aforementioned, the contemporary discussion of AWS is a classic debate between consequentialists (or instrumentalists) and deontologists, between technological optimists and pessimists. Our discussion is at once wider and narrower than both lines of argument. It is wider since it places the challenges emanating from AWS as part of the general phenomenon of exercise of state power in relation to the individual. It is narrower in the sense that it proposes to view the problem from an administrative-discretion point of view, which avoids the key pitfalls of the instrumental argument and the inconveniences of the deontological approach. Before proceeding, we briefly demonstrate the circular nature of the current debate.

A. *The Instrumentalist Discourse*

The instrumentalist debate regarding AWS is generally focused on their direct results in relation to the protection of individuals; on the problem of assigning responsibility for their ‘decisions’ *ex post*; and on their potential to diminish restraints on the resort to force by lowering the price of warfare in terms of human lives.²⁴ While the issues of responsibility and *jus ad bellum* are of much significance, we focus here on the more primordial issue of protection of victims of war.

The instrumental discourse is essentially an argument about optimistic or pessimistic perceptions of technology, coupled with optimistic or pessimistic perceptions of human nature. At the heart of the pessimist objections is deep skepticism regarding the ability of computers to *ever* achieve a level of sophistication required to apply the complex legal standards prevalent in modern

²⁴ For a useful summary of such objections see generally Autonomous Weapons Report.

IHL.²⁵ While some concede – although this in itself doubtful –²⁶ that AWS might sufficiently distinguish between civilians, combatants, and persons *hors de combat* during ‘traditional’ armed conflicts, virtually all pessimists argue that it is impossible to imagine a system capable of achieving good results when applying vaguer norms.

If placing the pessimistic approach in jurisprudential terms, the claim is that while it is possible to envision computers applying *rules*, it is hard-pressed to imagine them applying *standards*, which *per se* do not aim to predict the right legal outcome in any given situation.²⁷ The distinction between rules and standards is helpful, as it allows us to explain better common intuitions as to why some computerized processes seem acceptable while others seem objectionable – even in the context of ‘life and death’ decisions. For this reason, for instance, most people do not find it especially problematic that computers are delegated the power to calculate artillery trajectories²⁸ – which essentially amounts to applying a set of rules derived from geography and weather combined with safety-parameter rules; while it seems harder to accept computers making decision regarding direct participation in hostilities²⁹ or *jus in bello* proportionality,³⁰ precisely because such decisions are evaluated in light of legal standards rather than rules. This is especially true regarding proportionality, which essentially requires an ethical decision between the value of a military advantage and the expected civilian harm. It has been labeled a profoundly human qualitative

²⁵ See, e.g., Sharkey, ‘Evitability’, 788 –790.

²⁶ See N. Sharkey, ‘Cassandra or False Prophet of Doom: AI Robots and War’, *IEEE Intelligent Systems*, 23 (2008), 14 –17; Asaro, ‘On Banning Autonomous Weapon Systems’, 697.

²⁷ See Louis Kaplow, ‘Rules Versus Standards: An Economic Analysis’, *Duke Law Journal*, 42 (1992), 557 –629. On rules and standards in the context of IHL see Amichai Cohen, ‘Rules and Standards in the Application of International Humanitarian Law’, *Israel Law Review* 41 (2008), 4 – 67; see also Asaro, ‘On Banning Autonomous Weapon Systems’, (distinguishing between computers’ ability to apply ‘chess’ type rules versus legal norms that require ‘interpretation’).

²⁸ See, e.g., <http://udcusa.com/product-development-technical-services/indirect-fire-direction/>.

²⁹ The determination of direct participation in hostilities is shrouded in controversy and in any case complex. Compare N. Melzer, ICRC, *Interpretive Guidance on the Notion of Direct Participation in Hostilities Under International Humanitarian Law*, (2009) With K. Watkin, ‘Opportunity Lost: Organized Armed Groups and the ICRC “Direct Participation in Hostilities Interpretive Guidance’’, *NYU Journal of International Law and Politics*, 42 (2010), 641–695; According to some, determination of direct participation is especially problematic when AWS are concerned, since it requires the ability to assess intentions, which autonomous weapons are highly unlikely to have. See Asaro, ‘On Banning Autonomous Weapon Systems’, 697 –698.

³⁰ AWS Report, ¶¶70–73; Asaro, ‘On Banning Autonomous Weapon Systems’, 697; Human Rights Watch, ‘ 30 –

decision,³¹ not least because it is judged in terms of ‘common sense’ and the ‘reasonable military commander,’ which are classic legal standards that seem incompatible with technical autonomy.³² Naturally, these problems are exasperated in the ‘fog of war,’ where unforeseen factual combinations can fall outside the pre-programmed ‘design parameters’ of AWS, and the latter – lacking human ‘common sense’ – will be unable to adapt.³³ The pessimism regarding technology is usually accompanied with pessimism regarding programmers: even if complex ethical and legal standards could be embedded in computer software, they are likely to be distorted by programmers, at best because they are not themselves knowledgeable in the intricacies of the field.³⁴

While we sympathize with these concerns, the problem with the instrumental objections to AWS is that they can always be countered, at least analytically, by resort to optimistic hypotheticals regarding future technologies. Such arguments assume that the development of artificial intelligence capable of delivering good results is possible, likely or even inevitable.³⁵ Not only is this development possible – but it could also be, as the arguments go, a positive one. Indeed, like the instrumental objections to AWS, the optimists also rely on

³¹ Sharkey, *Evitability*, at 789–790; Asaro, ‘On Banning Autonomous Weapon Systems’, at 700 (‘applying IHL requires multiple levels of interpretation in order to be effective in a given situation’).

³² *Autonomous Weapons Report*, ¶¶70–73; Human Rights Watch, ‘Losing Humanity’, 32–33.

³³ See Asaro, ‘On Banning Autonomous Weapon Systems’, 691–692 (and sources cited therein).

³⁴ Keats Citron, ‘Technological Due Process’, 1261.

³⁵ See, e.g., K. and M. Waxman, ‘Law and Ethics for Robot Soldiers’, *Policy Review* 176 (2012), <http://www.hoover.org/publications/policy-review/article/135336> (suggesting that even though it is debatable whether AWS could ever be advanced enough to satisfy the ethical and legal principles of distinction and proportionality, the possibility should not be ruled out when discussing the normative approach towards autonomous weapons); Schmitt and Thurnher, ‘Out of the Loop’, 239–240, 262; see also R.C. Arkin, *Governing Lethal Behavior in Autonomous Robots* (CRC Press, 2009), pp. 65–66 (suggesting the theoretical possibility that the most complex standards of IHL – and even utilitarian or Kantian ethics – can be applied by artificial intelligence software through the employment of what he calls an ‘ethical governor.’); J.O. McGinnis, ‘Accelerating AI’, *Northwestern University Law Review*, 104 (2010), 1253–1269, 1256–1260; compare Asaro, ‘On Banning Autonomous Weapon Systems’, 699 (‘The empirical question is whether a computer, machine, or automated process could make each of these decisions of life and death and achieve some performance that is deemed acceptable. But the moral question is whether a computer, machine or automated process ought to make these decisions of life and death at all. Unless we can prove in principle that a machine should not make such decisions, we are left to wonder if or when some clever programmers might be able to devise a computer system that can do these things’).

consequentialist reasoning, focusing on the possible *results* of the deployment of AWS.

Thus, instrumental supporters of AWS point out the fact that if done right, autonomous technology will be much better than humans in alleviating death and suffering at war.³⁶ Indeed, excluding human beings from the battle ground will diminish the need for ‘force protection,’ which might in turn lower collateral damage inflicted on civilians.³⁷ In any case, the circular debate goes on and on. Confronting the optimistic assumptions regarding future technology, pessimists claim that relying on technological hypotheticals, should not inform our normative debate today.³⁸

On the legal level, Schmitt and Thurnher offer a robust instrumental defense of AWS, arguing that there is nothing *per se* unlawful in the process of autonomous targeting, assuming that their deployment and results correspond with IHL. As they claim, it is possible to envision algorithms that apply the principle of distinction, and are flexible enough to reasonably perform complex proportionality calculations.³⁹ Emphasizing outcomes, they remind us that humans, too, are less than perfect in applying IHL’s difficult standards.⁴⁰

However, they too concede that in the near future there is no choice but leaving the ‘proportionality decision’ – and other complex determinations – to humans. This human involvement, however, can be manifested in the mere decision to deploy the system in specific circumstances.⁴¹ Indeed, much of Schmitt and Thurnher’s reasoning shifts back the point of assessment to the decision to deploy the weapon.⁴² Needless to say, this claim assumes that the

³⁶ See, e.g., McGinnis, ‘Accelerating AI’, 1265–1267.

³⁷ *Ibid.*, 1266.

³⁸ Asaro, ‘On Banning Autonomous Weapon Systems’, 703 (describing the hypothetical technology argument as ‘rhetorical strategy’); see also Human Rights Watch, ‘Losing Humanity’, 27–28; Sharkey, ‘Evitability’, n.4.

³⁹ Schmitt and Thurnher, ‘Out of the Loop’, 243–250 (arguing, *inter alia*, that an autonomous weapon system is only unlawful *per se* if there are no circumstances in which it can be used discriminately); 250–257 (discussing the prospects of autonomous systems capable of adhering to the principles of distinction and proportionality); 262–265 (discussing cases of doubt).

⁴⁰ *Ibid.*, 257.

⁴¹ *Ibid.*, 257, 267.

⁴² For instance, if a weapon that cannot distinguish between combatants and civilians is deployed in a mixed environment, the deploying commander could commit the crime of indiscriminate attack. *Ibid.*, 278.

deploying officer can fully anticipate the system's actions, which is in itself a debatable contention.⁴³

In our eyes, the discussion of AWS should neither be guided by optimistic prophecies regarding the immediate results that can be achieved by new technologies, nor by unrelenting pessimism: indeed, some humility is called for when attempting to predict future technologies.⁴⁴ Therefore, a substantive discussion of AWS must transcend the circular and speculative discourse regarding their ability to deliver end results.⁴⁵

The instrumental arguments concerning AWS talk beyond each other not only with regard to future technology, but also concerning the putative effects of 'human nature' on battlefield outcomes. Interestingly, instrumentalists that support AWS are as pessimistic regarding human conduct as they are optimistic about technology's power; while those that are pessimistic regarding technology romanticize human behavior on the battlefield. Thus, the optimists claim that humanity will be better off with AWS, since the latter would be immune to negative aspects of the human psyche such as tendency for revenge, prejudice or sadism,⁴⁶ as well as to natural fear and panic that can result in material risk to civilians.⁴⁷ As phrased by McGinnis, it is 'mistaken' to assume that artificial intelligence would 'worsen, rather than temper, human malevolence.'⁴⁸ However, pessimists are quick to retort that the elimination of the human factor from the field does not only neutralize *adverse* human traits, but also positive ones such as common sense or compassion, which might allow humans to act mercifully even

⁴³ Some argue that due to the sheer complexity of modern computing 'no individual can predict the effect of a given command with absolute uncertainty.' See P. Lin et al, 'Autonomous Military Robotics' Report Prepared for the US Department of Navy, Office of Naval Research (2008), ¶6 (also stating that the notion that robot behavior is fully predictable as 'a common misconception'). Thus, their acts can become unpredictable in chaotic situations such as armed conflicts. See Autonomous Weapons Report, ¶42.

⁴⁴ See W. Wulf, 'Observations on Science and Technology Trends', in A.G.K. Solomon (ed.), *Technology Futures and Global Power, Wealth and Conflict* (CSIS, 2005)10, 10 –16. As candidly admitted by Noel Sharkey – one of the key critics of autonomous weapons – the possibility of a 'black swan event' in which new technology will allow autonomous weapons to apply complex standards well enough cannot be theoretically precluded. Sharkey, 'Evitability', note 4.

⁴⁵ Asaro, 'On Banning Autonomous Weapon Systems', 699.

⁴⁶ AWS Report, ¶54

⁴⁷ *Ibid.*, ¶70.

⁴⁸ McGinnis, 'Accelerating AI', 1254; see also Schmitt and Thurnher, 'Out of the Loop', 249.

when they could lawfully kill.⁴⁹ Paradoxically, both pessimists and optimists disregard the fact that human nature – whether good or bad – is also behind the development of technology itself, which can therefore (potentially) either reflect humanity’s benevolence or its malevolence.

It becomes clear that as always, assumptions about human nature are mutually-offsetting and are likely to remain inconclusive,⁵⁰ and thus cannot carry the debate forward. Thus, any attempt to decide whether the elimination of negative human traits from the battlefield is instrumentally positive, in terms of immediate results, is speculative and prone to arbitrariness.

B. The Deontological Discourse

The second level of the debate on AWS is concerned with deontological ethics. Arguments of this order posit that even if AWS could overcome the instrumental concerns discussed above – a possibility concerning which deontologists are of course pessimistic – and could therefore, in terms of results, perform well in relation to the values that IHL sets out to protect – there is *still* something inherently wrong in subjecting humans to machine-made ‘decisions’ to use lethal force.

At large, the deontological arguments are variations of two basic claims. The first has to do with the nature of war, and specifically the expectation that it would involve some element of mutual risk: as the argument goes, in order to distinguish it from one-sided killing, war must involve some measure of self-risk.⁵¹ According to this view, killing in war is not immoral, if war is understood as a series of individual self-defense actions undertaken by combatants against

⁴⁹ Human Rights Watch, ‘Losing Humanity’, 37–39

⁵⁰ See, e.g., N. Chomsky and M. Foucault, *Human Nature: Justice vs. Power – The Chomsky-Foucault Debate*, new edn, (The New Press, 2006); Peter Loftson, *Theories of Human Nature*, (Broadview Press, 1995), pp. 19–29. Arguments (and counter-arguments) regarding ‘human nature’ have accompanied the discourse of international law at least since Hobbes, Machiavelli and Grotius, the latter claiming, contra the former, that human beings are naturally inclined ‘to follow the direction of a well-tempered judgment, being neither led astray by fear or the allurements of immediate pleasure, nor carried away by rash impulse.’ Hugo Grotius, *On the Laws of War and Peace*, Vol. 2, (Clarendon Press, 1925), p.13; see also H. Lauterpacht, ‘The Grotian Tradition in International Law’, *British Yearbook of International Law*, 23 (1946), 1–53, 24–26.

⁵¹ AWS, ¶60.

each other.⁵² Killing through AWS is therefore unjustified unilateral killing.⁵³ However, this objection does not capture the salient dilemma of AWS: this is because it could equally apply also to *other* methods and tactics of warfare such as drones, cruise missiles and high-altitude bombing – and in the past, to crossbows, rifles, and artillery.⁵⁴ However, nobody seriously suggests, nowadays, a complete, deontological ban on the latter. This must mean that there is something additionally vexing in the concept of AWS.

The second type of deontological arguments concern the *nature* of the mechanized decision-maker on the one hand, and the human dignity of potential victims on the other. Arguments of this order point out the distance that AWS create between human beings and the consequences of their actions. As phrased by Special Rapporteur Christof Heyns, AWS create a ‘new dimension’ in this distance: not only will humans be removed physically from the battlefield – as drones already allow – but they also become detached *mentally* from decisions to kill.⁵⁵ Precisely because of this reason, for instance, the International Committee for Robot Arms Control (ICRAC) asserts that ‘machines should not be allowed to make the decision to kill people.’⁵⁶

From the outset, it is worthwhile to recall that as always, deontologists are placed in an awkward position when confronted by extreme hypotheticals. Again, instrumentalists are likely to point out that if the ultimate objective of law in armed conflict is to protect civilians, and assuming that AWS could be actually better than humans in doing so, the deontologist has three main (and unattractive) escapes: either to deny the plausibility of the hypothetical; or to withdraw, in some form, from her absolute position (threshold deontologism); or – and in our context, it is indeed a fitting phrase – to ‘bite the bullet’ and accept that greater

⁵² P.W. Kahn, ‘The Paradox of Riskless Warfare’, *Philosophy and Public Policy Quarterly*, 22 (2002), 2–8.

⁵³ It should be noted that such arguments can have an instrumental version as well: Kahn, for instance, argues that the immorality of riskless warfare will ultimately increase the adversaries’ attacks against the civilian population, for lack of other options. See Kahn, ‘The Paradox’, 2, 6.

⁵⁴ It was in this context that the argument was initially made. See, e.g., *ibid.*

⁵⁵ AWS Report, ¶¶26–27.

⁵⁶ Mission Statement of the International Committee for Robot Arms Control (2009), <http://icrac.net/statements/>.

loss of life is preferable to lesser loss of life, only because a machine is involved in the process.⁵⁷

The argumentation of Peter Asaro, a founding member of ICRAC, represents a sophisticated deontological position, based on considerations emanating from IHL and supplemented from ethics. Asaro's basic claim revolves around the principle that 'the authority to decide to initiate the use of lethal force cannot be legitimately delegated to an automated process.'⁵⁸ He grounds these notions on positive IHL, claiming that the latter 'explicitly requires combatants to reflexively consider the implications of their actions, and to apply compassion and judgement in an explicit appeal to their humanity.'⁵⁹ As he argues, what matters under IHL – for instance, in the context of proportionality – is not necessarily the 'specific calculation' but rather the ultimate 'deliberate consideration' undertaken by humans before using lethal force.⁶⁰ Asaro thus asserts that 'justice cannot be automated:' when computer processes replace human agents, due process is fundamentally compromised, as human judgment is constitutive of any system of justice.⁶¹

This reasoning brings us closer to a satisfying discussion of the dilemma of AWS, as it alludes to an understanding of warfare as an executive action. However, it still requires us to make a broader theoretical claim about the relations between warfare and public law. Absent such a claim, it is unclear to what extent IHL alone would require such a 'residual' possibility of human discretion, assuming an autonomous system that otherwise delivers reasonable results. Namely, Asaro's approach does not provide a sufficient theoretical link between the norms regulating armed conflict – in which hostilities are conducted against an adversary, and concepts of procedural justice or fairness such as 'due process' usually reserved to domestic law. In short, arguments of this order tell us

⁵⁷ See L.A. and M. Moore, 'Deontological Ethics', in E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (2012), <http://www.science.uva.nl/~seop/entries/ethics-deontological/#WeaDeoThe>.

⁵⁸ Asaro, 'On Banning Autonomous Weapon Systems', 689.

⁵⁹ *Ibid.*, 700.

⁶⁰ *Ibid.*

⁶¹ *Ibid.*, 700–701. This is because according to Asaro, the essence of due process is 'the right to question the rules and appropriateness of their application in a given circumstance, and to make an appeal to informed human rationality and understanding.' *Ibid.*, 700.

why AWS might jeopardize concepts of procedural justice such as due process; but they do not tell us why these are relevant, *as such*, in situations of warfare.

In sum, as we have demonstrated above, neither the instrumental, nor the deontological discourses – in their various forms – are entirely satisfying. In the next Part, we attempt to offer a third prism through which to analyze the complex question of AWS. In particular, this prism can serve as a theoretical bridge between warfare and administrative concepts of justice and due process.

III. REFRAMING THE DEBATE: WARFARE AS AN EXERCISE OF ADMINISTRATIVE POWER

A. *War as Governance: Warfare as an Exercise of Administrative Power*

As mentioned above, some of the deontological arguments about AWS refer to ideas of procedural justice and due process. However, their shortcoming lies precisely in that they fail to demonstrate how, or to what extent, this framework is relevant to the law regulating wars. In order to argue that notions such as due process apply during war, we thus have to also make an argument about war itself. As we argue, war under contemporary international law must be understood as a form of governance⁶² – a fact that spawns, as we demonstrate, administrative-legal obligations.

Justice and due process seem foreign to the traditional idea of war. Indeed, in the past, the laws of war were quite explicit in their obliviousness to these as we understand them in domestic law.⁶³ For instance, the 1863 Lieber Code bluntly stated that ‘[t]he citizen or native of a hostile country is thus an enemy, as one of the constituents of the hostile state or nation, and as such is subjected to the

⁶² Compare B. Kingsbury et al, ‘The Emergence of Global Administrative Law’, *Law and Contemporary Problems*, 68 (2005), 15 – 61, 17 (identifying administrative action/governance on the international level as all ‘decisions that are neither treaty-making nor simple dispute settlements between parties.’); see also E. Benvenisti and A. Cohen, ‘War is Governance: Explaining the Logic of the Laws of War from a Principal-Agent Perspective’, *Michigan Law Review* 112 (2013), 1363– 1415; E. Lieblich, ‘Show Us The Films: Transparency, National Security and Disclosure of Information Collected by Advanced Weapon Systems Under International Law’, *Israel Law Review*, 45 (2012), 459 –491, 483.

⁶³ We set aside the idea that war itself is a form of ‘punishment’ – and thus of justice-making – found in early just war thought. See D. Luban, ‘War as Punishment’, *Philosophy and Public Affairs*, 39 (2012) 299 –330.

hardships of the war.’⁶⁴ Binding individuals together as a political collective, essentially holding them liable for state-action, negates their individual agency; this is hardly an example of justice or due process as the terms are regularly understood. Indeed, war was not about justice, as it was generally perceived as ‘a contention between States for the purpose of overpowering each other,’ in which each party was entitled to defend its own interests aggressively while weakening the adversary.⁶⁵ Even nowadays, the difference between justice-making and war-fighting is a key difference between understandings of the law enforcement paradigm versus that of hostilities.⁶⁶ Thus, absent significant theoretical expansion, IHL is not commonly understood as a ‘system of justice,’ as implied by Asaro,⁶⁷ but rather a system aiming to minimize harm in a given, inherently unjust situation, where in any event, military necessity remains the prime consideration.⁶⁸ Naturally, the divergence between war, justice and due process is most acute where conflicts are international or transnational, in which a ‘social contract’ between the warring parties cannot be assumed. Nonetheless, even in cases of civil wars, it could be argued that in some cases, the social contract is at least temporarily broken, thereby modifying the normal relations within the state in a manner that affects the regular application of the principles of due process.

The traditional view of war, thus, usually entailed at least some assumption of collectivity: war was a ‘horizontal’ confrontation between collectives, in which the ‘enemy’ comprised both the state apparatus and

⁶⁴ General Orders No. 100: Instructions for the Government of the Armies of the United States in the Field (The Lieber Code), art. 21. It should be noted that this view was primarily dominant in Anglo-American thought. See L. Oppenheim, *International Law*, 2nd edn, Vol. 2, (Longman’s, Green and Co., 1912) §57.

⁶⁵ *Ibid.*, §58 (‘As war is a struggle for existence between States, no amount of individual suffering and misery can be regarded; the national existence and independence of the struggling State is a higher consideration than any individual well-being.’)

⁶⁶ See E. Lieblich, ‘Quasi-Hostile Acts: The Limits on Forcible Disruption Operations Under International Law’, *Boston University International Law Journal*, 32 (2014), 101–155, 110–117.

⁶⁷ Asaro, ‘On Banning Autonomous Weapon Systems’, 701. The question whether there is a system of justice beyond the state is of course strongly contested among philosophers. See T. Nagel, ‘The Problem of Global Justice’, *Philosophy and Public Affairs* 33 (2005), 113–147; J. Cohen and C. Sabel, ‘Extra Rempublicam Nulla Justitia?’, *Philosophy and Public Affairs*, 34 (2006), 147–175.

⁶⁸ As understood by the mainstream of writers, IHL is ‘intended to minimize human suffering without undermining the effectiveness of military operations.’ Y. Dinstein, *The Conduct of Hostilities Under the Law of International Armed Conflicts*, (Cambridge University Press, 2004), p. 17.

citizenry.⁶⁹ Accordingly, it was possible to dismiss offhand any special obligations between a state and the enemy's citizens. Each state was responsible, beyond certain minimal obligations, only to safeguard the rights and interests of its own citizens against the belligerent actions of the other.⁷⁰ War thus envisioned a 'vertical' system of exclusive protection between sovereign and citizenry. These vertical obligations were only transferred from one sovereign (e.g. the defender) to another (e.g. the attacker) when the latter occupied its territory.⁷¹ It is easy to see why, when holding this view of war, claims that AWS violate due process can be shrugged as out of touch.

This perception of war was perhaps sustainable in an age dominated by conflicts between functioning states. Nowadays, conversely, most wars are fought between states and non-state actors, and armed violence is limited neither spatially nor temporally.⁷² In many cases – and virtually in all instances where AWS are likely to be deployed – armed force will be used by advanced militaries against non-state actors, in failing states or territories.

In such conflicts, the absence of functioning sovereigns where hostilities are conducted results in a gap in civilian protection. This gap is a product of three main characteristics of asymmetric conflicts. *First*, in many cases, such conflicts involve significant intermingling between militants and civilians, which naturally increases civilian harm. This intermingling is caused both by deliberate failure, in many cases, of armed groups to distinguish themselves from the civilian population, and also since guerilla tactics, almost always, involves at least some

⁶⁹ See Janina Dill, 'Should International Law Ensure the Moral Acceptability of War', *Leiden Journal of International Law*, 26 (2013), 253–270, 259. However, it should be noted that the collective view of war was already challenged by Rousseau. See J. Rousseau, *The Social Contract*, (1762), Book 1, Ch. IV. Rousseau's view about war became dominant in Continental Europe during the 19th century, but was not influential in British and American thought. In any case, as noted by Oppenheim – who unequivocally supported the collective view – this dispute did not have substantial legal consequences in international law. Oppenheim, *International Law*, §57.

⁷⁰ See, e.g., W. Hays Parks, 'Air War and the Law of War', *Air Force Law Review*, 32 (1990), 1–225, 21 (noting that traditionally, 'collateral civilian casualties ... were not regarded as the responsibility of an attacker, as the ability to limit such casualties lay more with the defender or, for that matter, with the civilian population itself.').

⁷¹ Hague Convention (IV): Respecting the Laws and Customs of War on Land, The Hague, 18 October 1907, Regulations, Arts. 42–43 (hereinafter Hague Regulations).

⁷² See Uppsala Conflict Data Program (UCDP), 'Armed Conflict Dataset' v.4-2014, 1946–2013, http://www.pcr.uu.se/research/ucdp/datasets/ucdp_prio_armed_conflict_dataset/.

concealment that puts civilians at risks.⁷³ *Second*, in many cases, armed groups are unaccountable, whether normatively, because of shortcomings of the legal regime;⁷⁴ whether in practice, due to lack of motivation to comply;⁷⁵ or because of absence of pressure from constituencies, as such actors are not susceptible to political pressure as sovereigns in functioning states are. When armed groups are unaccountable, significant risk can be transferred to civilians under their control. *Third*, in asymmetric conflicts, the traditional concept of ‘victory’ is virtually non-existent.⁷⁶ In such scenarios states might rephrase the notion of victory in ambiguous terms. One possibility is to discuss victory in terms of winning the sympathy of the local population, and therefore seek to reduce civilian harm.⁷⁷ However, the opposite process is just as plausible: states engaged in what they perceive as a zero-sum conflict might do just the opposite, and adopt ‘deterrence’ as their main objective, making indirect use of incidental harm to civilians to further that purpose.⁷⁸ In sum, since in most situations of non-international armed conflicts individuals lack any substantial effective institutional protection, a protection gap emerges; if humanitarian considerations are to be taken seriously,

⁷³ See, for instance, the use of underground tunnels by Hamas in the 2014 Israel-Hamas conflict. Frequently, the tunnels’ openings are inside civilian buildings. See Harriet Sherwood, ‘Inside the Tunnels Hamas Built: Israel’s Struggle Against New Tactic in Gaza War’, *The Guardian* (2 August, 2014), <http://www.theguardian.com/world/2014/aug/02/tunnels-hamas-israel-struggle-gaza-war>.

⁷⁴ See generally Lisbeth Zegveld, *Accountability of Armed Opposition Groups in International Law*, (Cambridge University Press, 2002).

⁷⁵ See, e.g., D. Richemond-Barak, ‘Nonstate Actors in Armed Conflicts: Issues of Distinction and Reciprocity’, in W.C. Banks (ed.), *New Battlefields, Old Laws: Critical Debates on Asymmetric Warfare* (Columbia University Press, 2011) pp. 106–129, p. 107

⁷⁶ As successful confrontations with non-state actors usually require ‘long term change’ rather than tactical advantages. See G. Blum, ‘The Fog of Victory’, *European Journal of International Law*, 24 (2013), 391–421, 391–394 (2013).

⁷⁷ *Ibid.*, 408; US Department of the Army, *The US Army/Marine Corps Counterinsurgency Field Manual*, (University of Chicago Press, 2007), p. 294.

⁷⁸ Compare Giora Eiland, ‘The Third Lebanon War: Target Lebanon’, INSS Strategic Assessment 11, no. 2 (2008). Of course, deterrence is a vague and speculative advantage. In essence, it can serve to justify almost any harm, and thus cannot comport with legitimate military goals both under *jus ad bellum* and *jus in bello*. The law on the use of force, as generally understood, restricts self-defense to what is necessary to counter a specific threat. See Tom Ruys, *‘Armed Attack’ and Article 51 of the UN Charter: Evolutions in Customary Law and Practice*, (Cambridge University Press, 2010), p. 94; IHL defines military necessity narrowly, as comprising only as the need to ‘weaken the military forces of the enemy.’ Declaration Renouncing the Use, in Time of War, of Explosive Projectiles Under 400 Grammes Weight, St. Petersburg, 1868. In the same vein, military advantage that can be measured in relation to incidental civilian harm must be ‘concrete and direct.’ Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts (Protocol I), 8 June 1977, Art. 51(5)(b) (hereinafter API).

this gap must result in some ‘diagonal’ responsibilities of protection between the potentially affected individuals and the attacker who is in a position to decide their fate.⁷⁹

The above analysis suggests that we must view modern warfare as a form of exercise of state power vis-à-vis individuals, rather than a horizontal engagement between sovereigns. It therefore becomes possible to understand it as closer to the exercise of administrative, executive action than to traditional war.⁸⁰ Once it is perceived this way, warfare should be subjected to widely accepted notions of accountability and responsiveness that are reflected in domestic administrative law that governs executive decision-making. Of course, administrative power (and its scrutiny) is responsive to circumstances. When we suggest that battlefield decisions can be analyzed as executive actions we do not imply that the full spectrum of administrative law applies to each targeting decision. What we refer to are basic principles.⁸¹

Granted, traditionally, administrative duties during warfare were triggered only once a territory was occupied.⁸² However, it is important to stress that ultimately, the ‘active agent’ generating such obligations is *control* rather than this or that legal status. Indeed, in recent years, the concept of control has been significantly broadened to include numerous situations in which state agents affect

⁷⁹ See generally Eyal Benvenisti, ‘Rethinking the Divide between Jus ad Bellum and Jus in Bello in Warfare Against Nonstate Actors’, *Yale Journal of International Law*, 34 (2009), 541–548. When discussing AWS in terms of executive action, it becomes clearly revealed that questions relating to computerized decision-making extend well beyond the issue of targeting. For instance, during occupation, but also in the ‘invasion’ stage, autonomous weapons can be utilized for all kinds of measures of control, some bordering law enforcement activities. Theoretically, they can be used for curfew enforcement, for dispersing demonstrations, population control or to destroy property. These possibilities are worthy of further exploration, but are beyond the scope of this paper. For AWS in law enforcement see Christof Heyns, ‘Human Rights and the Use of Increasingly Autonomous Weapons Systems During Law Enforcement’ (forthcoming).

⁸⁰ Compare Kingsbury et al, ‘The Emergence of Global Administrative Law’, 17 (asserting and explaining that nowadays ‘much of global governance can be understood and analyzed administrative action.’).

⁸¹ We are aware of the claim that discussing contemporary warfare in administrative terms might have adverse by-effects of normalizing a neverending state of war (and thank Itamar Mann for pointing out this problem). However, problems of this order are a consistent risk when engaging in any form of transnational regulation, and thus we cannot resolve them here. For a discussion of these dynamics in another context see I. Mann, ‘Dialectic of Transnationalism: Unauthorized Migration and Human Rights’, 1993–2013, *Harvard Journal of International Law*, 54 (2013), 315–391. Specifically with regard to ‘new’ wars compare P.W. Kahn, ‘Imagining Warfare’, *European Journal of International Law*, 24 (2013), 199–226; with S. Moyn, ‘Drones and Imagination: A Response to Paul Kahn’, *European Journal of International Law*, 24 (2013), 227–233.

⁸² See Hague Regulations, Arts. 42–43.

individuals even when control over territory is not complete.⁸³ States have been expected to act as benevolent administrators even without boots on the ground, at least in relation to the *functions* they control vis-à-vis the civilian population.⁸⁴ In our context, when making a targeting decision, a state agent is making a life and death decision regarding an individual. As such, the process of targeting represents a form of control *par excellence*, even if not always accepted as such in positive international law concerning accumulation of jurisdiction.⁸⁵

However, we do not have to adopt a radical perception of control in order to ground the existence of administrative-like duties during warfare. This is since complementing the expanding notion of control is an evolving perception of sovereignty itself. As we suggest, this transformation is from a view of sovereignty as a ‘trusteeship’ of certain body-politic, to trusteeship of *humanity* at large. The concept of sovereignty as trusteeship of humanity, its justifications – as well as its possible critiques – were expounded elsewhere.⁸⁶ Here we shall only outline its basic (and relevant) aspects.

The gist of this approach is that states should bear a residual obligation to take other-regarding considerations in their decision-making processes; and that in fact, this idea is increasingly reflected in international law.⁸⁷ The appeal of this approach lies in ever-growing interdependence. This interdependence means that sovereigns’ decisions concerning development, natural resources or – in fact – any type of regulation, no longer affect only the polity of the territorial state, but also lives of individuals in other states: power and its effects are no longer aligned.⁸⁸ Put simply, the perception of sovereignty as a global trust rests upon the notion

⁸³ *Al Skeini v. UK*, App. No. 55721/07, ECtHR (Judgment, 7 July 2011), ¶¶131–140; Human Rights Committee, ‘General Comment 31, Nature of the General Legal Obligation on States Parties to the Covenant’, UN Doc. CCPR/C/21/Rev.1/Add.13 (2004), ¶10.

⁸⁴ This has been notably the case regarding the relationship between Israel and Gaza. See, e.g., Gisha – Legal Center for Freedom of Movement, ‘Scale of Control: Israel’s Continued Responsibility in the Gaza Strip’, 14 November, 2011; Orna Ben-Naftali et al, ‘Legal Opinion Concerning Supply of Electricity and Water to the Gaza Strip’, 20 July, 2014, <http://gisha.org/UserFiles/File/publications/letters/letter-en-20-7-14.pdf>; in other contexts, see *Al-Skeini v. UK*, Concurring Opinion of Judge Bonello, ¶¶11–12.

⁸⁵ See *Bankovic v. Belgium*. App. No. 52207/99, ECtHR (Decision, 12 December 2001), ¶¶62–82.

⁸⁶ E. Benvenisti, ‘Sovereigns as Trustees of Humanity: On the Accountability of States to Foreign Stakeholders’, *American Journal of International Law*, 107 (2013), 295–333; see also ‘AJIL Symposium: Sovereigns as Trustees of Humanity’, *Opinio Juris* (23 July, 2013), <http://opiniojuris.org/2013/07/24/ajil-symposium-sovereigns-as-trustees-of-humanity/>.

⁸⁷ Benvenisti, ‘Sovereigns as Trustees’, 297.

⁸⁸ *Ibid.*, 298.

that those who exercise power bear responsibility, even if the results affect those found beyond borders. It is, in its essence, a corollary of the idea of the *equal moral worth* of all, but also can be justified on perceptions of global welfare and justice.⁸⁹

However, from considerations both normative as well as practical, global trusteeship does not call for the eradication of the state in favor of a cosmopolitan utopia of world-federation, nor does it require that states relinquish their self-interest altogether. It merely recognizes a pragmatic administrative-legal notion, according to which a sovereign's decision-making process must take effects over others seriously.⁹⁰ Importantly, since this responsibility is inherent in the concept of sovereignty, it applies to *every* exercise of sovereign power, independent of specific treaty obligations.⁹¹ Thus, global trusteeship spawns several minimal obligations which are firmly rooted in traditional notions of administrative law. Relevant to our context are the obligation to take others' interests into account (including human dignity), which implies a duty to exercise discretion when sovereign actions bear cross-border effects.⁹²

Thus, the mere fact that a certain armed conflict crosses national borders does not, in itself, negate diagonal-administrative responsibilities, even when traditional effective control is absent. How could this understanding affect the debate on AWS? The answer lies in applying the principle of equal moral worth. In this context, Margalit and Walzer offer a helpful argument. Using an illuminating thought-exercise, they argue that when armed forces operate, they must assume the same level of risk – when choosing means and methods of warfare – whether operating in the vicinity of the enemy's civilians or among their own.⁹³ We do not claim that states cannot prefer their own civilians across the board. As David Luban convincingly notes, the issue is not whether states can

⁸⁹ *Ibid.*, 301–313.

⁹⁰ *Ibid.*, 300–301.

⁹¹ *Ibid.*, 300.

⁹² Benvenisti, 'Sovereigns as Trustees', 313– 325.

⁹³ A. Margalit and M. Walzer, 'Israel: Civilians & Combatants', *NY Review of Books* (14 May, 2009). In their example, a state's village is taken by the enemy. Will the state be morally justified to act differently, in terms of the risk undertaken by its forces and the resulting collateral damage to civilians, according to the citizenship of the persons within that village? Their conclusion is a firm no. Compare Asa Kasher & Amos Yadlin, *Military Ethics of Fighting Terror: an Israeli Perspective* 43 *J. MIL. ETHICS* 3 (2005).

require soldiers to take ‘heroic risks’ in order to save their ‘own’ civilians (they can), but rather that a state cannot treat ‘enemy’ civilians beneath minimum acceptable standards, in a manner that it would never treat its own.⁹⁴ The delineation of these minimum standards cannot be exhausted here. A possible cut-off point, suggested by Luban, refers to the identity of the risk creator: states have to treat ‘enemy’ civilians as their own, in terms of alleviating risk, *at least* when the risk is created by the state’s own violence – for instance, when a targeting decision is made. In other words, as Luban stresses, while it is possible to defend the claim that a state is only obliged to defend its *own* citizens – but not others – against threats by an external enemy, when the risk is created by its own actions (and to the extent that this is the case, we should add), the state must give the lives of the endangered civilians on the other side the same weight that it would give its own.⁹⁵ If we apply Margalit, Walzer and Luban’s logic to the question of AWS, we may ask states that consider deploying such weapons, *whether they would be willing to use them in situations where their own citizenry could be affected by ‘decisions’ made by such systems.*

It seems that some states have already answered this question negatively, albeit in an implied manner, by imposing limitations on the computerization of executive decision-making. For instance, Article 15 of EU Directive 95/46/EC,⁹⁶ enshrines the right of every person ‘not to be subject to a decision which produces legal effects concerning him ... which is based solely on automated processing of data’ The Article’s key rationale is to limit predetermined decision-making, which, as we shall see, is a key problem emanating from AWS.⁹⁷ Granted, Article 15 allows automated processes if they are ‘authorized by law which also lays

⁹⁴ See D. Luban, ‘Risk Taking and Force Protection’, Georgetown Public Law & Legal Theory Research Paper No. 11-72, (2011), 12, 46.

⁹⁵ Luban, ‘Risk Taking’ 32 –33 (‘Of course the soldier’s job exists to protect co-civilians, not foreign civilians, against their enemies... But it would completely beg the question to assert that the soldier’s obligation to protect civilians against the soldier’s own violence runs only to fellow citizens’); but see I. Porat and Z. Bohrer, ‘Preferring One’s Own Civilians: May Soldiers Endanger Enemy Civilians More than They Would Endanger Their State’s Civilians?’, *George Washington International Law Review* (Forthcoming, 2014). We set aside the discussion whether it could be said, in any context, that the “risk” is created, in the true sense, by the attacking party.

⁹⁶ EU Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the Protection of Individuals with Regard to the Processing of Personal Data and on the Free Movement of Such Data, Art. 15(1).

⁹⁷ *Ibid.*, Arts. 3–6

down measures to safeguard the data subject's legitimate interests.⁹⁸ However, in order not to render the Article meaningless, this would imply *at least* some form of appeal, to ensure that the final decision is not computerized.⁹⁹ In the jargon of AWS, this would require maintaining a 'human in the loop' before making irreversible decisions. Admittedly, in recent years, 'automated decision-making systems' have gained some ground in the US, but these are used to apply *rules* rather than *standards*, and in any case are not involved in life and death decisions.¹⁰⁰ Indeed, if states limit computerized decisions with regard to their own citizens – and in matters far less acute than life and death in the battlefield – it is questionable whether it would be legitimate to subject *others* to such decisions. The potential effect of an administrative perception of warfare over the question of AWS is clear: it means that AWS would be subject to additional, *residual* constraints, *even* if such weapons would perform reasonably, in terms of *immediate results*, under IHL.

It should be noted, that in actuality – and although not phrased as such – basic notions of administrative law are not entirely foreign to modern, positive IHL. For instance, the requirement to take '*constant care*' in the conduct of military operations,¹⁰¹ is reminiscent of the traditional administrative-law notion of the duty to exercise discretion when making decisions, upon which we elaborated shortly. The duty to give effective warning 'unless circumstances do not permit';¹⁰² or the obligation to direct an attack against the least dangerous target, in terms of civilian harm, when the attack of several military objectives would spawn the same military advantage, also requires active, administrative-like discretion.¹⁰³ The rights of internees to elect representatives to treat with the detaining power,¹⁰⁴ are not unlike well recognized administrative obligations of due process and fair-hearing. The gradual elimination of belligerent reprisals

⁹⁸ *Ibid.*, Art. 15(2)(b).

⁹⁹ Keats Citron, 'Technological Due Process', 1265 (exemplifying automatic administrative systems that allow for appeals).

¹⁰⁰ *Ibid.* 1252, 1263 (providing examples for such decision-making systems in the US).

¹⁰¹ API, Art. 57(1).

¹⁰² *Ibid.*, Art. 57(2)(c).

¹⁰³ *Ibid.*, Art. 57(3).

¹⁰⁴ Convention (IV) Relative to the Protection of Civilian Persons in Time of War, Geneva, 12 August 1949, Art. 102.

against persons *hors de combat*,¹⁰⁵ demonstrates, too, a shift from a perception of warfare as a series of reciprocal blows between black-boxed sovereigns, into a more complex phenomenon, in which some diagonal responsibility is established between sovereigns and the adversary's civilians.

The same administrative law approach can be derived from IHRL, which is also considered applicable during armed conflict and – in some cases, extraterritorially.¹⁰⁶ IHRL requires that the limitation or deprivation of rights, including the right to life, be subject to due process in the course of limiting those rights.¹⁰⁷ As the ECtHR asserted, ‘the decision-making process leading to measures of interference must be fair and such as to afford due respect to the interests of the individual.’¹⁰⁸ And the British House of Lords, in interpretation the requirement under the ECHR that limitations on rights be ‘in accordance with the law’ equated ‘arbitrariness’ with compliance with the rule of law by public officials.¹⁰⁹ Whereas in *McCann v. UK*,¹¹⁰ the ECtHR emphasized the *ex-post* duty to conduct an effective investigation when individuals had been killed as a result of the use of force by state agents, as derived from the duty to protect the right to life,¹¹¹ there is little doubt that the *ex-ante* duty to exercise discretion diligently and thoroughly also flows from the same obligation. The administrative law perception of warfare can thus provide the additional conceptual framework for understanding what constitutes ‘arbitrary’ deprivation of life during armed

¹⁰⁵ See R. Kolb, ‘The Main Epochs of Modern International Humanitarian Law Since 1864 and their Related Dominant Legal Constructions’, in K. Mujezinovi Larsen et al (eds.), *Searching for a ‘Principle of Humanity’ in International Humanitarian Law*. (Cambridge University Press, 2012), pp. 23–71, pp. 52–53.

¹⁰⁶ For an overview see N. Lubell, ‘Human Rights Obligations During Military Occupation’, *International Review of the Red Cross*. 94 (2012), 317–337, 317–323 (2012).

¹⁰⁷ See E. Benvenisti, *The Law of Global Governance*, (Hague Academy of International Law, 2014) Ch. 3.

¹⁰⁸ *Taşkin v. Turk.*, App. No. 46117/99, (2005), ¶¶ 124–125.

¹⁰⁹ *R (on the application of Gillan (FC) and another (FC)) v Commissioner of Police for the Metropolis and another*, [2006] UKHL 12, ¶ 34. See also *Gillan & Quinton v. U.K.*, Appl. No. 4158/05, ¶¶ 76–77 (2010) (“well established case-law [indicates that] the words ‘in accordance with the law’ require the impugned measure both to have some basis in domestic law and to be compatible with the rule of law.” This means that “legal discretion granted to the executive [cannot] be expressed in terms of an unfettered power. Consequently, the law must indicate with sufficient clarity the scope of any such discretion conferred on the competent authorities and the manner of its exercise”).

¹¹⁰ *McCann v. U.K.*, App. No. 18984/91 (1996).

¹¹¹ *Ibid.*, ¶ 161. The right to life under IHRL is enshrined, for instance, in Art. 6(1) of the International Covenant on Civil and Political Rights, 16 December 1966.

conflict. At large, the administrative law point of view can reveal that exercise of proper administrative discretion is indeed a *precondition* for the application of human rights standards, considering that the latter are almost always vague and balanced versus other rights and interests.¹¹²

The Martens Clause can also be invoked in order to introduce the administrative law approach into IHL. The Martens Clause, as introduced in the 1899 and 1907 Hague Conventions, and reiterated in Additional Protocol I, famously recognizes the ‘laws of humanity’ and the dictates of public conscience as residual gap fillers when positive IHL seems silent.¹¹³ Indeed, relying on this clause, some deontological arguments claim that AWS inherently contradict the principle of humanity and therefore must be banned.¹¹⁴ While the Martens Clause is undoubtedly significant, basing an objection to AWS strictly on principles of humanity or public conscience is problematic because it begs the question of what is the nature of the relations between those deploying AWS and individuals on the ‘other side.’ In other words, simply invoking the Martens Clause without further elaboration seems to overstretch it.¹¹⁵ But informed by the trusteeship approach, this gap can be bridged to add support and content to the reliance on the Martens Clause as an additional ground for the administrative perception.

B. Autonomous Weapons and the Problem of Pre-bound Discretion

As noted above, positive IHL requires belligerents to take ‘constant care’ to spare civilians.¹¹⁶ ‘Constant care’ is an admittedly vague standard, and as such open to various interpretations, lenient as well as constraining. Precisely here, the

¹¹² For a related argument see M. Koskenniemi, ‘The Future of Statehood’, *Harvard International Law Journal*, 32 (1991), 397–410, 399 (1991). Indeed, the claim that discretion is a *precondition* for the fulfillment human rights could preempt, for instance, the validity of a treaty which would formulate mandatory algorithms for autonomous weapons.

¹¹³ Convention (II) with Respect to the Laws and Customs of War on Land, The Hague, 29 July 1899; Hague (IV); API, Art. 1(2); see generally T. Meron, ‘The Martens Clause, Principles of Humanity, and Dictates of Public Conscience’, *American Journal of International Law*, 94 (2000), 78–89.

¹¹⁴ Human Rights Watch, ‘Losing Humanity’, 35–36.

¹¹⁵ See Meron, ‘The Martens Clause’, (‘... the Martens clause does not allow one to build castles of sand. Except in extreme cases, its references to principles of humanity and dictates of public conscience cannot, alone, delegitimize weapons and methods of war, especially in contested cases’); see also Schmitt and Thurnher, ‘Out of the Loop’, 275–276; for a discussion see also T. D. Evans, ‘Note: At War With Robots: Autonomous Weapons Systems and the Martens Clause’, *Hofstra Law Review*, 41 (2013), 697–733.

¹¹⁶ API, Art. 57(1).

administrative law perception can guide us, as it allows us to understand properly this notion as a requirement to exercise *continuous discretion* when conducting hostilities.¹¹⁷

The obligation to exercised discretion imposes upon the administrative authority the duty to consider, within the confines of its legal authority, each decision to exercise power in light of the specific goals of the norm that the executive is bound to promote, including the relevant rights and interests affected in the case at hand. This obligation calls for a duty to *constantly* exercise discretion. Of course, this duty implies a prohibition – and indeed the invalidity – on fettering one’s discretion. The authorizing agency must be willing to listen to ‘anyone with something new to say’ and to alter or waive its policies in appropriate cases.¹¹⁸ The very idea of delegating decision-making authority to actors is to have them exercise their discretion in individual cases given the specific circumstances of each case. If there was no need to pay attention to the specific circumstances and to allow for fresh thinking, the superior organ could have made the decision itself. While some pre-commitment by administrative agencies is indeed a legitimate tool to promote transparency and equal treatment, it seeks to stipulate the boundaries of discretion, not to *negate* it altogether; moreover, pre-commitment must be of such nature that be altered in real-time if circumstances require.¹¹⁹

Indeed, the fundamental requirement that discretion cannot be fettered is also reflected in general doctrines of international law, through the concept of ‘reasonableness’ (or in IHL language, of ‘feasibility’).¹²⁰ The mere idea of reasonableness *must* require leaving open the possibility to make adjustments in one’s policies, through the exercise of discretion. For instance, the international law of non-navigational use of watercourses requires upstream states to make ‘equitable and reasonable’ use of trans-boundary water sources.¹²¹ It seems

¹¹⁷ Compare Schmitt & Thurnher, ‘Out of the Loop’, 259–260.

¹¹⁸ See *British Oxygen v. Minister of Technology*, [1971] AC 610, HL (UK).

¹¹⁹ See E. Magill, ‘Agency Self-Regulation’, *George Washington Law Review*, 77 (2009), 101–144, 104, 117 (2009).

¹²⁰ See, e.g., O. Corten, ‘The Notion of ‘Reasonable’ in International Law: Legal Discourse, Reason and Contradictions’, *International and Comparative Law Quarterly*, 48 (1999), 613–625.

¹²¹ Convention on the Law of the Non-Navigational Uses of International Watercourses, New York, 21 May 1997, Art. 5(1).

obvious that a state's attempt to predetermine what would be 'reasonable' use in all circumstances, and to act upon such criteria without reviewing them according to changing circumstances, will be a clear violation of the reasonableness principle.¹²² If there could be a clear a-priori definition of what is to be considered 'reasonable', the parties would have agreed on it in advance.

There are two key rationales for the obligation not to fetter one's discretion. The first emanates from the rights of the affected individual. The executive is bound to give 'due respect' to individuals, by considering the effects of a specific act on individuals, in light of prevailing circumstances.¹²³ This is an essential characteristic of the trust relations which form the basis of administrative power, one which is especially significant in the context of hazardous activities, such as warfare.¹²⁴

The second justification for the duty to exercise discretion has to do with decision-making quality: we assume that in the long run, 'good' executive decisions cannot be taken, in a complex world, without making constant adjustments.¹²⁵ These adjustments, which require constant discretion, are necessary due to epistemological human limitations, which Hart identified as comprising 'relative ignorance of fact' and 'relative indeterminacy of aim.'¹²⁶ These 'handicaps' limit any attempt to regulate decision-making in advance.¹²⁷

During hostilities, the duty to exercise discretion requires the active, ongoing intention not to inflict harm on civilians.¹²⁸ The components of this duty require the military commander to exercise discretion both when planning the attack,¹²⁹ and, importantly, *during* attacks, up to the last moment before pulling the trigger.¹³⁰ Naturally, the last requirement applies primarily to 'those

¹²² *Ibid.*, Art. 6 (requiring to take into account 'all relevant factors and circumstances' when utilizing international watercourses).

¹²³ Benvenisti, 'Sovereigns as Trustees', 314.

¹²⁴ Compare *ibid.*, 316.

¹²⁵ See, e.g., Barth and Arnold, 'Artificial Intelligence and Administrative Discretion', 338, 348–349.

¹²⁶ H.L.A. Hart, 'Discretion', *Harvard Law Review*, 127 (2013), 652–665, 661–664.

¹²⁷ *Ibid.*

¹²⁸ See M. Walzer, 'Coda: Can the Good Guys Win', *European Journal of International Law*, 24 (2013), 433–444, 437.

¹²⁹ API, Art. 57(2)(a)(i)

¹³⁰ *Ibid.*, Art. 57(2)(b).

executing' the attack.¹³¹ In our context, in an out-of-the-loop scenario, the duty to exercise discretion 'in the last moment' would have to be performed by the AWS.

This outcome is problematic when taking into account proper administrative discretion. AWS, as aforementioned, are only capable of exercising *technical* autonomy; they cannot engage in the metacognition required in order to exercise 'true' discretion in real-time. Therefore, their use reflects a case where executive discretion is stringently bound in advance – through the pre-programmed algorithms that govern their 'behaviour'. The deployment of AWS thus runs counter to the two rationales justifying the prohibition on fettering one's administrative discretion – respect for the individual and decision-making quality. First, pre-binding targeting decisions, for instance, clearly contradicts the notion of 'due respect' for the individual, since the potentially harmed individual is not considered at all in real-time, but is rather 'factored' in pre-determined processes. Second, AWS cannot be reconciled with the inherent need to make constant adjustments in a complex world, considering the epistemological limitations mentioned above. Indeed, if the legislator could not have promulgated an order to shoot in advance, if state parties could not have agreed in advance whether a certain military act was lawful or not, how could a military commander, at any level, fetter her discretion or of anyone else under her command?¹³²

In this context, it is important to reiterate that the duty to exercise discretion is concerned with *long-term* decision-making quality. Accordingly, it should not be judged in light of isolated results: while this or that pre-determined decision can bring about good results, any administrative system based on rigid

¹³¹ C. Pilloud et al (eds.), Commentary on the Additional Protocols of 8 June 1977 to the Geneva Conventions of 12 August 1949, (Martinus Nijhoff Publishers, 1987) p. 686.

¹³² A question related to the limitation on fettering discretion concerns *delegation* of authority. In general, administrative legal doctrines differentiate between 'lean' delegation of technical powers on the one hand, and delegation of discretionary power on the other. The latter, naturally, is placed under stricter scrutiny. See. e.g., J. Freeman, 'Private Parties, Public Functions and the New Administrative Law', *Administrative Law Review*, 52 (2000) 813–858, 823 (noting the difference between delegation of 'management' and delegation of 'rule-making' authority). The use of semi-AWS – for instance, those that 'lock' on targets and track them – could reflect the former type of delegation. Fully autonomous ones, conversely, are reminiscent of the latter. The use of such weapons is, perhaps, a form of 'double' delegation: first, by delegating to programmers the power to design the weapon's 'discretion,' it reflects a delegation of rule-making power; and then, by delegating the power to make battlefield decisions to a non-human, it reflects a delegation of powers to make individual decisions. Compare Keats Citron, 'Technological Due Process', 1253 (noting that computer programs 'seamlessly combine rulemaking and individual adjudications without the critical procedural protections owed [in] either of them.').

rules, taken as a whole, is likely to diminish the overall quality of decision-making, or worse – can potentially bring about catastrophic results. The latter problem can be easily demonstrated in the context of autonomous targeting. For instance, while a human administrator (or commander) is prone to make an occasional ‘wrong’ decision, her decision is unique and affects the specific case at hand. Even if she is in the higher organizational (or command) echelons, her decision can still (at least in theory) be corrected by judgment exercised by subordinates. Arguably, the diversity of human discretion is mutually correcting, ultimately leading to reasonable overall results.¹³³ Conversely, decisions based on ‘wrong’ algorithms, or, for that matter, on any pre-determined rigid rule, will be repeated by all systems in which they are embedded, to potential catastrophic effects.¹³⁴ While it is relevant to ask whether the life-saving potential of ‘good’ AWS is not sufficient to outweigh such concerns, one should not lose sight of the catastrophic potentials of a systemic mistake.

In a sense, computerized pre-commitment constitutes a significant shift in the ‘vertical landscape’ of discretion. If IHL requires common sense throughout the command chain – from the general all the way down to the private, the so-called ‘discretion’ embedded in AWS would probably represent mainly the pre-programmed perceptions of higher echelons. As such, it will negate field decision-making which is not only a legal requirement under IHL, but can also be carried out better.¹³⁵ Naturally, the importance of on the field decision-making (sometimes referred to as ‘street-level’ decision-making) increases in intensively ‘discretionary’ environments, which involve complex decision-making under uncertainty.¹³⁶ Most modern armed conflicts clearly constitute such environments.

¹³³ See, in another context. C.R. Sunstein, *Infotopia: How Many Minds Produce Knowledge*, (Oxford University Press 2006).

¹³⁴ See, e.g., J. Frémont, ‘Computerized Administrative Decision Making and Fundamental Rights’, *Osgoode Hall Law Journal*, 32 (1994), 818–831, 826–827 (1995); the failures of automated administrative systems in the US exemplify this problem. See Keats Citron, ‘Technological Due Process’, 1256, 1267–1271. The same effect can emanate from programmers’ bias. *Ibid.* 1262.

¹³⁵ *Ibid.* 1262–1263

¹³⁶ E.Z. Brodtkin, ‘Street-Level Research: Policy at the Front Lines’, in M.C. Lennon and T. Corbett (eds.), *Policy Into Action: Implementation Research and Welfare Reform*, (The Urban Institute Press, 2003), pp. 145–164, p. 145, 151 (arguing that policy, in discretionary environments, should be analyzed ‘at the point of delivery’).

A major pitfall in the instrumentalist arguments that defend AWS can be found precisely regarding the problem of discretion. In order to reassure us that a human still remains ‘in the loop’ – meaning, that discretion is indeed exercised – some instrumentalists claim that the only salience of AWS is that human discretion takes place on a different *temporal* level. As they argue, human discretion is embedded into the system through the human discretion of the programmers. As phrased by Schmitt and Thurnher, ‘in an attack involving autonomous weapon systems ... targeting decisions remain subjective and continue to be made by humans. What may differ is *the phase in the targeting process when the subjective determinations occur.*’¹³⁷ As Sharkey critically describes this process, it is ‘like writing a knitting pattern or recipe.’¹³⁸ To put it in our terms, what Schmitt and Thurnher suggest (and Sharkey criticizes) is the binding of discretion *ex ante*. However, this is *precisely* what makes AWS problematic, when analyzed through the lens of proper administrative action.

Granted, Schmitt and Thurnher reassure us that the ‘decisive juncture’ of human discretion is found at the point where the decision is made to deploy the weapon in specific circumstances.¹³⁹ However, this reassurance is not entirely convincing, for two reasons. *First*, it relies on the contestable assumption that commanders would be able to foresee how weapons would ‘act,’ at least in a manner sufficient to amount to the taking of ‘constant care’ to spare civilians. Here lies a paradox: for the weapon’s action to be predictable, it must be built to follow simple, rigid rules, which seems to mean that it would not be capable to perform IHL’s complex standards to begin with. However, if it is unpredictable – meaning, that it attempts to apply complex standards in a changing environment – then deploying it *per se* violates the duty to take *constant* care. Moreover, since humans tend to attribute to computers ‘far more intelligence than they actually possess,’ and are therefore unlikely to ‘override’ their decisions,¹⁴⁰ it is possible to envision a phenomenon of ‘reliance-creep,’ where the duty to take constant care will be ‘delegated’ more and more to the computer’s pre-programmed discretion.

¹³⁷ Schmitt and Thurnher, ‘Out of the Loop’, 267.

¹³⁸ Sharkey, ‘Evitability’, 789.

¹³⁹ Schmitt and Thurnher, ‘Out of the Loop’, 267.

¹⁴⁰ Barth and Arnold, ‘Artificial Intelligence and Administrative Discretion’, 348; Keats Citron, ‘Technological Due Process’, 1254.

In this context, it is also unclear to what extent allowing autonomous systems to select and ‘recommend’ human targets would alleviate the problem of discretion, considering the ‘automation bias’ that may occur when human beings assess computerized recommendations.¹⁴¹

Second, even under Schmitt and Thurnher’s approach, the fact remains that the *last* decision to use force will be made by the weapon, upon pre-bound discretion. Indeed, Schmitt and Thurnher suggest that systems would include methods of ‘adjustment,’ allowing commanders to ‘dial up and down’ values in order to accommodate specific circumstances.¹⁴² If these adjustments can be made during all stages of the attacks, including up to the last minute – this might indeed solve the problem of pre-bound discretion. But then, it would be fair to ask whether such weapons can be considered autonomous to begin with.

In sum, the reading of relevant legal norms in light of administrative law notions, and specifically the duty to take constant care and thereby exercise discretion, results in the conclusion that technically autonomous weapon systems cannot be allowed to make final targeting decisions.

C. Subjecting Warfare to Administrative Law Standards and the Problem of Pre-Bound Discretion: Possible Challenges

Indeed, the administrative law approach at large, as well as our emphasis on the problem of pre-bound discretion, can raise some challenges or objections. However, as we briefly demonstrate, they are not sufficient to negate either of these ideas.

One objection to the idea of residual ‘administrative’ constraints could be that during armed conflict, IHL constitutes the *lex specialis*, replacing, as such, ‘peace time’ notions of administrative law. This argument, essentially, mirrors the same debate regarding the application of international human rights law during armed conflict.¹⁴³ In this context, we subscribe to the widely accepted view that

¹⁴¹ See *ibid.*, ‘Technological Due Process’, 1272.

¹⁴² Schmitt and Thurnher, ‘Out of the Loop’, 267.

¹⁴³ For the treatment of this question by the ICJ see *Legality of the Threat or Use of Nuclear Weapons*, Advisory Opinion 8 July 1996, ICJ 226 ¶25; *Legal Consequences of a Wall in the Occupied Palestinian Territory*, Advisory Opinion 9 July 2004, ICJ 136, ¶106; *Case Concerning*

even if specific IHL norms can serve as *lex specialis* in certain circumstances and thus override obligations from other areas of law, IHL *as such* does not *replace* other normative systems as a whole; rather, there is significant space for complementarity between norms belonging to different systems or for interpreting IHL to include basic notions of administrative law concerning the proper exercise of discretion.¹⁴⁴ Thus, IHL and administrative law constraints can exist side by side. In this sense, it is interesting to note that administrative law is actually a more ‘convenient’ partner for dual-application with IHL than IHRL, *at least* in light of a conservative understanding of the latter. This is because much of administrative law is concerned with secondary norms of *process*: it chiefly aims to regulate the manner in which decisions are made,¹⁴⁵ rather than to replace the legal system’s primary norms.¹⁴⁶ Indeed, precisely because of its focus on process, the administrative approach allows us to discuss AWS critically beyond the usual instrumental-deontological circle: it departs from the act-utilitarianism espoused by some proponents of AWS, since it can serve to constrain specific actions even if their immediate *results* seems just (or legal); but it is also not deontological since it is minded towards achieving long-term results in terms of the quality of decision-making.¹⁴⁷

Armed Activities on the Territory of the Congo (DRC v. Uganda), Judgment of December 19 2008, ICJ, ¶215.

¹⁴⁴ For a discussion see, e.g., M. Milanović, ‘Norm Conflicts, International Humanitarian Law, and Human Rights Law’, in O. Ben-Naftali (ed.), *International Humanitarian Law and International Human Rights Law*, (Oxford University Press, 2011), pp. 95 – 128; N.R. Quéniwet, ‘Introduction: The History of the Relationship Between International Humanitarian Law and Human Rights Law’, in R. Arnold and N. Quéniwet (eds.), *International Humanitarian Law and Human Rights Law: Towards a New Merger in International Law*, (Brill, 2008), pp. 1–14, pp. 9–11.

¹⁴⁵ See Kingsbury et al, ‘The Emergence of Global Administrative Law’, 15 (defining global administrative law as ‘as comprising the mechanisms, principles, practices, and supporting social understandings that promote or otherwise affect the accountability of global administrative bodies, in particular by ensuring they meet adequate standards of transparency, participation, reasoned decision, and legality, and by providing effective review of the rules and decisions they make.’); see also *ibid.*, 29.

¹⁴⁶ Perhaps this is why the Supreme Court of Israel found no special problem to rule that principles of Israeli administrative law apply extraterritorially to military action in the West Bank, while being far more reluctant to concede that IHRL applies to such operations. For the historical reluctance to apply IHRL in the occupied territories see O. Ben-Naftali and Y. Shany, ‘(2008)’, *Israel Law Review*, 37 (2004), 17–118 For application of administrative standards of proportionality in targeting by the Israeli Supreme Court see, e.g., HCJ 769/02, *The Public Committee against Torture in Israel v. The Government of Israel*, 62 (1) PD 507, ¶ 40 [2006] (Isr.)

¹⁴⁷ It could be argued that our approach is one of rule-consequentialism: it looks at the consequences promoted by a decision making process guided by certain rules. See B. Hooker,

Challenges can also be presented to our claim that AWS present a special problem of pre-bound discretion. One possible objection could be that regular, ‘dumb’ weapons, such as artillery rounds or rockets, also reflect pre-bound discretion: indeed, once fired – upon the discretion of the shooter – there is no turning back. However, we reject this *ad absurdum* argument on two counts. *First*, the time gap between the exercise of human discretion and the weapons’ impact, in the case of regular kinetic weapons, is usually negligible. Any change of circumstances between the last human decision and impact, such that would require re-engaging of discretion is unlikely (although not entirely impossible). This is not the case when deploying AWS, which by nature are expected to act in a relatively prolonged manner.¹⁴⁸ *Second*, a ‘dumb’ weapon does not presume to execute complex legal standards itself, but rather follows simple physical rules. To begin with, these weapons do not attempt to make distinction or proportionality decisions. The dumbness of the artillery shell must be taken into account by its launcher, who is also responsible for the potentially unlawful outcomes. In contrast, the sophistication of the AWS nullifies its launcher’s *opportunity* to exercise discretion.

A related claim could be that nonetheless, the essence of military discipline is that human soldiers, too, do not exercise discretion, but rather follow orders. As such, the argument goes, even today, warfare assumes fettering the discretion of soldiers through their subjection to superiors. However, military thinkers have long been aware that in practice, soldiers *do* exercise discretion, for better or for worse, even when receiving clear orders.¹⁴⁹ Normatively, it is clear that international law requires discretion even from the lowest-ranked soldier, *inter alia* through the negation of ‘superior orders’ defense in certain circumstances.¹⁵⁰

‘Rule Consequentialism’, in E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, (2011), <http://plato.stanford.edu/entries/consequentialism-rule/>.

¹⁴⁸ Similar problems are raised by anti-personnel landmines. See generally M.A. Lawson et al (eds.), *To Walk Without Fear: The Global Movement to Ban Landmines* (Oxford University Press, 1998).

¹⁴⁹ See, e.g., K. Kagan, *The Eye of Command*, (The University of Michigan Press, 2006), p. 176 (discussing Clausewitz).

¹⁵⁰ For an overview see E. van Sliedregt, *Individual Criminal Responsibility in International Law*, (Oxford University Press, 2012), pp. 287–306. Z. Bohrer, ‘The Superior Orders Defense: A

D. Diagonal Obligations towards Combatants, Pre-Bound Discretion and the Shrinking Window of Surrender

As we have demonstrated, the administrative law perception of warfare spawns diagonal obligations towards the adversary's civilians. Do comparable obligations exist in relation to enemy *combatants*? In this context, traditional laws of war set forth down a bright-line rule: combatants, as well as civilians that directly participate in hostilities could be attacked and killed, with no need to assess their individual threat: the 'necessity' of their targeting was presumed, unless they were *hors de combat*: meaning, they have surrendered, been captured or injured.¹⁵¹ In accordance with the bright-line rule, the laws of war as traditionally understood do not require attackers to grant enemy combatants the chance to surrender prior to staging the attack: as opposed to law enforcement actions, under the hostilities paradigm, the assumption is that enemy troops represent a threat, and thus it is presumed that they have not surrendered, absent a clear expression to that effect.¹⁵² Indeed, this remains the prevailing view.

However, recent years have seen an increasingly vibrant debate, challenging the enduring validity of this bright-line rule. Generally, the debate has 'migrated' from the question of targeted killings, and specifically the duty to arrest, if possible, civilians directly participating in hostilities,¹⁵³ to encompass also the rights of soldiers in international armed conflicts – whether through a reinterpretation of the concept of military necessity¹⁵⁴ or by an expansive understanding of the notion of *hors de combat*.¹⁵⁵ Whatever the basis for putative

Principal-Agent Analysis', *Georgia Journal of International and Comparative Law*, 41 (2012), 1–74.

¹⁵¹ Lieber Code, Art. 15 ('Military necessity admits of all direct destruction of life or limb of armed enemies.');

G. Blum, 'The Dispensable Lives of Soldiers', *Journal of Legal Analysis*, 2 (2010), 115–170, 123–126 (2010).

¹⁵² See J.M. Henckaerts and L. Doswald Beck, ICRC, *Customary International Humanitarian Law*, vol. 1, (Cambridge University Press, 2005), Rule 47, p. 164.

¹⁵³ See Melzer, 'Interpretive Guidance', Pt. IX; *The Public Committee against Torture in Israel v. The Government of Israel*, ¶40; but see W. Hays Parks, 'Part IX of the ICRC 'Direct Participation in Hostilities' Study: No Mandate, No Expertise, and Legally Incorrect', *NYU Journal of International Law and Politics*, 42 (2010), 769–830, 783–785.

¹⁵⁴ See e.g. Blum, 'Dispensable Lives'.

¹⁵⁵ R. Goodman, 'The Power to Kill or Capture Enemy Combatants', *European Journal of International Law*, 24 (2013), 819–853; but see M.N. Schmitt, 'Wound, Capture or Kill: A Reply to Ryan Goodman's "The Power to Kill or Capture Enemy Combatants"', *European Journal of*

obligations to prefer capture, under some circumstances, of enemy combatants, it seems that their underlying assumption is that there are *diagonal* obligations between a state and enemy combatants *qua* individuals; and that such obligations might extend to scenarios beyond traditional scenarios of *hors de combat* (such as incapacitation or capture). Although these obligations have not yet crystalized as widely accepted *lex lata* – and additionally, it is clear that even if only from practical reasons, they are much weaker than obligations a state might owe to the enemy’s *civilians*¹⁵⁶ – the discourse that requires that state parties consider the *individuality* of enemy combatants opens the door for administrative law reasoning also in this context, at least as *lex ferenda*.

This would mean that the problem of AWS and pre-bound discretion can arise also when civilians are not at all endangered by their deployment. This realization allows us to assess the compelling argument presented by Schmitt and Thurnher, that even if AWS should not be deployed where civilians are put at risk, this in itself should not result in their total ban. As they argue, this is because just like concerning any other weapon, the valid question, when considering the legal banning of AWS, is whether their deployment would be illegal in *all* circumstances. As Schmitt and Thurnher point out, deploying AWS in battlefield scenarios where only combatants are present, cannot be *per se* unlawful.¹⁵⁷ However, as the debate on kill or capture demonstrates, it seems that targeting combatants, too, might require some discretion.

Diagonal state-combatant obligations might have implication in other contexts, *beyond* the problem of discretion. For instance, they can urge us to reconsider the seemingly well-established notion of ‘surrender’ – as consisting not only of a duty to afford quarter once a combatant has clearly surrendered, but, perhaps, also as an obligation to leave open or at least not excessively narrow the *option* to surrender. An illuminating analogy through which we can understand better this obligation is through the idea that those affected by administrative

International Law, 24 (2013), 855 –861; J.D. Ohlin, ‘The Duty to Capture’, *Minnesota Law Review*, 97 (2013), 1268 –1342.

¹⁵⁶ See, e.g., J. McMahan, ‘Gaza: Is Israel Fighting A Just War in Gaza’, *Prospect* (5 August, 2014), <http://www.prospectmagazine.co.uk/philosophy/gaza-is-israel-and-hamass-conflict-a-just-war> (arguing that at large, obligations towards civilians are more demanding than obligations towards soldiers).

¹⁵⁷ Schmitt and Thurnher, ‘Out of the Loop’, 249 –250.

decisions must be given the opportunity to be heard.¹⁵⁸ Of course, translating the right to be heard to the battlefield is not straightforward: it cannot be understood in the exact manner as the same obligation during peace. However, some parallels can still be made. For instance, the well-established prohibition on ordering that no quarter will be given¹⁵⁹ is essentially a duty to leave open – at least *passively* – the option of surrender, when clearly expressed. The option of surrender is thus akin to an open window for a primitive type of hearing: the surrendering combatant is letting the other side know that she has chosen her right to life over her right to participate in hostilities,¹⁶⁰ and therefore there are no longer grounds to exercise executive power against her.

Taking into account the objectives of IHL, it is safe to assume that increasing opportunities to surrender is a desired social good. This is so not only due to the individual human rights of troops, but also because maximizing the option to surrender minimizes the adversary's motivation to 'fight to the last drop.' The question thus arises whether AWS widen or narrow the window in which surrender is feasible. Importantly, this dilemma arises not only when AWS are deployed in the vicinity of civilians, but also in scenarios where civilians are not at all present.

It goes without saying that surrender is not always viable: it requires some physical proximity, and awareness of the enemy's presence.¹⁶¹ Thus, the obligation to allow surrender can by no means be an absolute *positive* duty under IHL – as such a demand will virtually replace the doctrine of targeting under the hostilities paradigm, with the use of force continuum entrenched in IHRL. However, bearing in mind the human rights of combatants, it is still worthwhile to consider the effect of technological advancements on the feasibility of the option to surrender. These have not been entirely linear. Interestingly, for centuries, the battlefield scenarios in which surrender remained a viable possibility have decreased in line with the development of technology. For instance, it was nearly always possible to surrender in a swordfight; it became harder to do so when fire

¹⁵⁸ Benvenisti, 'Sovereigns as Trustees', 318.

¹⁵⁹ Henckaerts and Doswald Beck, *Customary IHL*, Rule 46, pp. 161 –163 (detailing the development of the principle).

¹⁶⁰ See API, Art. 43(2).

¹⁶¹ See Henckaerts and Doswald Beck, *Customary IHL*, Rule 47, p. 168.

powder was introduced; and virtually impossible in the era of ballistic weapons and air warfare, at least in most circumstances.¹⁶² However, at least in some situations, the advent of smart technology can potentially reverse this trend, since weapons (and delivery platforms) can be controlled almost until final impact.¹⁶³

Assuming that AWS would be deployed as replacement for ground forces, the question arises whether they will increase or *decrease* surrender opportunities. Indeed, when a human is in control of the final decision to pull the trigger, there is a minimal time gap between discharging the weapon and impact: during this gap, of course, surrender is impossible. However, in the time gap between such decisions, there is (in theory) ample opportunity to surrender – even during the assault stage. Conversely, AWS – at least as currently envisioned – might narrow the ‘surrender window’ available to the adversary.¹⁶⁴ ‘Swarm technology’ is a prime example. In swarm attacks, large numbers of small, autonomous systems will engage the enemy, rapidly overwhelming it.¹⁶⁵ For example, as recently revealed by Lt. Gen. Benny Gantz, Israel’s Chief of General Staff, the IDF aims to deploy, in the future, ‘swarms of autonomous vehicles, robots and drones, at sea and maybe even land, interconnected, relying on micro and nano-technology.’¹⁶⁶ Thus, if more and more maneuvers will be conducted through swarm attacks, it seems that the surrender window will be narrowed.

Indeed, as aforementioned, the shrinking sphere of surrender is a constant corollary of technological developments. Nonetheless, the novelty of AWS, in this context, lies in a combination of three considerations: *first*, in contrast, say, to ballistic missiles (and perhaps a bit like landmines) AWS create a wider temporal

¹⁶² Compare D. Leigh, ‘Iraq War Logs: Apache Crew Killed Insurgents who Tried to Surrender’, *The Guardian* (22 October, 2010), <http://www.theguardian.com/world/2010/oct/22/iraq-war-logs-apache-insurgents-surrender>.

¹⁶³ A famous example is a 1991 incident where Iraqi soldiers have surrendered to an American drone. See P.W. Singer, ‘Military Robots and the Laws of War’, *The New Atlantis*, Winter (2009), 25–45, 28.

¹⁶⁴ Of course, it is theoretically possible that technology be developed that does not narrow the window of surrender. However, we find this unlikely, since one of the key interests behind states’ development of such weapons has to do with their quickness. In other words, narrowing the enemy’s options on the temporal level is one of the key rationales for their development to begin with. See Krishnan, ‘Killer Robots’, pp. 40–41. Furthermore, even if autonomous weapons be somehow designed as to not narrow the window of surrender, we will still run into the problem of pre-bound discretion.

¹⁶⁵ Schmitt and Thurnher, ‘Out of the Loop,’ 240.

¹⁶⁶ Address at BESA Center (9 October, 2013), <http://besacenter.org/new-at-the-besa-center/idf-chief-staff-benny-gantz-speaks-besa-center/>.

gap between deployment and use of lethal force, the same gap in which surrender is theoretically possible. Thus, they might further reduce the temporal window for surrender. *Second*, the lightning-quickness of such weapons, when engaging, will not only challenge human supervision, but will also all but incapacitate the adversary's possibility to submit. *Third*, AWS represent a technological leap. An administrative law perception of warfare implies that increased power spawns increased obligations. For instance, a state that possesses such advanced systems can be expected, by virtue of its power, *at least* not to further reduce the sphere of possible surrender, when technology (such as the availability of semi-autonomous weapons) allows otherwise.

CONCLUSION

The debate on AWS will surely intensify in the near future, in line with the advancement in technology. This Chapter sought to demonstrate that an analysis of the deployment of such weapons through the prism of administrative law might advance our understanding of the main challenges posed by such weapons. When looking at modern warfare as an executive action, notions of administrative law require that it would be conducted while exercising constant discretion. As we have demonstrated, AWS, at least when understood as technically autonomous, cannot exercise such discretion in the 'true' sense.

While the former is definitely true in cases where civilians might be at risk from such weapons, we have also advanced a preliminary argument regarding the deployment of AWS against *combatants*. As we claim, in line with the increasing calls to reconsider the traditional understanding of military necessity in relation to targeting combatants, the problem of discretion can arise also in this context. Furthermore, AWS might also *narrow* the window of surrender, which we identify as a problematic outcome, *at least* in terms of *de lege ferenda*.