

PROGRAMMED TO OBEY: THE LIMITS OF LAW AND THE DEBATE OVER MEANINGFUL HUMAN CONTROL OF AUTONOMOUS WEAPONS

Yuval Shany and Yahli Shereshevsky*

"What man wants is simply independent choice,
whatever that independence may cost and wherever it
may lead"¹

F. Dostoevsky

ABSTRACT

New military technologies are transforming the contemporary battlefield, raising complex ethical and legal questions previously unaddressed. This Article makes three novel contributions to the debate on Autonomous Weapon Systems (AWS) and military AI in the legal and ethical literature. First, it puts forward a normative argument against AWS—even if they outperform humans in adhering to the rules governing the conduct of hostilities. This argument is grounded in the critical importance of the human capacity to act over

* Yuval Shany is the Hersch Lauterpacht Chair in International Law, former Dean of the Law Faculty of the Hebrew University of Jerusalem, a former member of the UN Human Rights Committee (2013–2020), and an Accelerator Fellow, Oxford Ethics in AI Institute (2024–2025). Yahli Shereshevsky is an Associate Professor (Senior Lecturer), Faculty of Law, University of Haifa, and a Principal Investigator at the Minerva Center for the Rule of Law under Extreme Conditions. The authors would like to thank Tammy Harel-Ben Shahr, David Hughes, Markus Wagner, and the Participants in the European Society of International Law Workshop on Posthumanism, Technology, and Peace and Security, the Oxford Ethics in AI Work in Progress Seminar Series, the Three Generations of Digital Human Rights seminar and the "Artificial Intelligence and International Law" writers' workshop co-hosted by the Stockton Center for International Law, U.S. Naval War College, and the University of Reading for their excellent comments on earlier versions of this essay. The research was conducted with the support of ERC Grant No. 101054745: the Three Generations of Digital Human Rights (DigitalHRGeneration3), <https://3gdr.huji.ac.il/>.

1. FYODOR DOSTOEVSKY, NOTES FROM UNDERGROUND 24 (Richard Pevear & Larissa Volokhonsky trans., 1994).

and beyond the strict letter of the law. The Article contends that this capacity is central to the regulation of warfare, which permits, rather than obligates, the use of force against legitimate targets. Second, it offers a doctrinal analysis of International Humanitarian Law (IHL) and International Human Rights Law (IHRL)—the two principal legal regimes that regulate armed conflicts under international law—providing a fresh perspective on how they intersect in the context of AWS. Finally, the Article explores the extent to which its normative argument is persuasive in the context of military AI beyond AWS, an area that is rapidly evolving and already extensively employed in current conflicts. It examines the similarities and differences between these emerging technologies and reflects on their implications for the desirable regulation of both technologies.

TABLE OF CONTENTS

INTRODUCTION – FRAMING THE DISCUSSION	111
I. AUTONOMOUS WEAPONS AND MEANINGFUL HUMAN CONTROL – THE IMPORTANCE OF CHOICE	117
A. The debate over meaningful human control.....	117
B. IHL and normative pragmatism	119
C. Dignitary justifications for meaning human control.....	121
D. Meaningful human control and the importance of choice	124
E. Agency as a component of dignity	126
F. Meaningful human control and the structure of IHL	128
II. IHL, IHRL, LEX SPECIALIS AND MEANINGFUL HUMAN CONTROL	130
A. A doctrinal basis for a meaningful human control requirement under IHL	130
B. Legal risks associated with relying on general principles in IHL	132
C. A doctrinal basis for a meaningful human control requirement under IHRL.....	135
D. The relationship between IHL and IHRL in the Context of AWS	137
III. MEANINGFUL HUMAN CONTROL AND MILITARY AI BEYOND AWS	140
CONCLUSION	146

INTRODUCTION – FRAMING THE DISCUSSION

In this essay, we undertake the challenging task of writing something new in the continuous debate about autonomous weapons. While the vast majority of the literature focuses on concerns that such weapons will be incompatible with specific norms of International Humanitarian Law (IHL)², we suggest that their use, even in strict compliance with IHL, could make them fundamentally incompatible with the general humanitarian ethos underlying IHL.

From *THE TERMINATOR* (1984) to *Black Mirror's "Metalhead"* (2017) and *The Creator* (2023), popular culture has long depicted autonomous weapons as dystopian threats—cold, ruthless, relentless, and seemingly uncontrollable.³ These portrayals reflect a deep public anxiety about a future in which machines, rather than humans, decide who lives and dies in and around the battlefield. But what if reality is more nuanced than the movies? Proponents of Autonomous Weapon Systems (“AWS”) argue that such weapons could out-perform human-operated weapon systems and generate more humane outcomes, including lower collateral harm to civilians and less human suffering.⁴ For the purpose of this essay, we accept that AI-controlled

2. We use here the common term IHL to discuss the body of international law that regulates armed conflicts. Others sometimes refer to IHL as the Laws of War or the Laws of Armed Conflict (LOAC). See Emily Crawford, Annabelle Lukin & Jacqueline Mowbray, *The Terminology of the Law of Warfare*, 14 J. INT'L HUM. LEGAL STUD. 197, 197–222 (2023).

3. Many scholars have discussed the role of popular culture and technophobia in debates over new technologies in general and new military technologies in particular. See, e.g., Chaz Arnett, *Dystopian Dreams, Utopian Nightmares: AI and the Permanence of Racism*, 112 GEO. L.J. 1299 (2023); Steve Calandrillo & Nolan Kobuke Anderson, *Terrified by Technology: How Systemic Bias Distorts U.S. Legal and Regulatory Responses to Emerging Technology*, 2022 U. ILL. L. REV. 597 (2022); Yahli Shereshevsky, *Are All Soldiers Created Equal? – On the Equal Application of the Law to Enhanced Soldiers*, 61 VA. J. INT'L L. 271 (2020); Kevin L. Young & Charli Carpenter, *Does Science Fiction Affect Political Fact? Yes and No: A Survey Experiment on “Killer Robots,”* 62 INT'L STUD. Q. 562 (2018); Chris Jenks, *False Rubicons, Moral Panic, & Conceptual Cul-de-Sacs: Critiquing & Reframing the Call to Ban Lethal Autonomous Weapons*, 44 PEPP. L. REV. 1 (2016); Rebecca Crootof, *The Killer Robots Are Here: Legal and Policy Implications*, 36 CARDOZO L. REV. 1837 (2014).

4. We do not take a stand in this ongoing debate over the ability of AWS to out-perform humans, since our argument is relevant to both cases. See Lena Trabucco & Kevin Jon Heller, *Beyond the Ban: Comparing the Ability of Autonomous Weapon Systems and Human Soldiers to Comply with IHL*, 46 FLETCHER F. WORLD AFFS. 15 (2022) (reviewing the current debate over AWS and suggesting that it tends to downplay human limitations).

weapon systems may be more precise than human-controlled weapon systems. However, we suggest that even with high levels of precision and a superior ability to follow the law, there might be strong reasons to object to the deployment of AWS—including that they systematically apply force to the full extent legally authorized.

AWS are weapon systems that “can select and engage targets without further human intervention.”⁵ A paradigmatic example of AWS is a drone loitering over a battlefield, using an operational algorithm to identify its targets, then firing a missile at those targets. The only human involvement in the targeting operation is the initial design of the system and the decision to deploy it to a specific theatre of war or military operation.⁶ The same AI technology enables the deployment of comparable ground- and naval-based AWS.⁷ An even more advanced generation of AWS involves an across-domain combination of threat identification, selection of suitable means and methods of response, and activation of autonomous weapon systems—all without human intervention.⁸

In the context of international law governing armed conflicts, the legality of AWS has become a focal point of academic discussion.⁹

5. Ingvild Bode & Hendrik Huelss, *Autonomous Weapons Systems and Changing Norms in International Relations*, 44 REV. INT’L STUD. 393, 394 (2018). Note that there is still no formal definition in international law for AWS (or of comparable terms used in the public discourse, such as Lethal Autonomous Weapon Systems (LAWS), or “Killer Robots”).

6. Euisun Hwang, *Lethal Autonomous Weapons: The Next Frontier in International Security and Arms Control*, STAN. INT’L POL’Y REV. (Jan. 30, 2025) (describing that AWS systems “can independently identify and engage targets without human intervention” according to “robust algorithms”), <https://fsi.stanford.edu/sipr/content/lethal-autonomous-weapons-next-frontier-international-security-and-arms-control>.

7. John Yoo, *Embracing the Machines: Rationalist War and New Weapons Technologies*, 105 CALIF. L. REV. 443, 452–54 (2017).

8. A prototype system of this nature—the Joint All-Domain Command and Control System—has already been put into use by the U.S. military. See U.S. DEP’T OF DEF., SUMMARY OF THE JOINT ALL-DOMAIN COMMAND AND CONTROL (JADC2) STRATEGY (Mar. 2022).

9. Although a comprehensive list is beyond the scope of this paper, numerous academic papers discuss AWS. See, e.g., Kevin Jon Heller, *The Concept of “the Human” in the Critique of Autonomous Weapons*, 15 HARV. NAT’L SEC. J. 1 (2023); Christopher M. Ford, *Autonomous Weapons and International Law*, 69 S.C. L. REV. 413 (2017); Rebecca Crootof, *War Torts: Accountability for Autonomous Weapons*, 164 U. PA. L. REV. 1347 (2016); Tim McFarland & Tim McCormack, *Mind the Gap: Can Developers of Autonomous Weapons Systems Be Liable for War Crimes?*, 90 INT’L L. STUD. 361 (2014); Jack M. Beard, *Autonomous Weapons and Human Responsibilities*, 45 GEO. J. INT’L L. 617 (2014); Michael N.

Before briefly introducing this Article's contributions to that discussion, we wish to frame our analysis, explicitly stating what this Article seeks and does not seek to do.

One segment of the ongoing debate asks whether existing international law norms—primarily the IHL and international human rights law (IHRL)—can be applied to new technologies (i.e., evolution), or if the emergence of new technologies requires a new legal framework (i.e., revolution).¹⁰ While some states actors contend that the current legal framework can effectively regulate AWS,¹¹

Schmitt & Jeffrey S. Thurnher, “*Out of the Loop*”: *Autonomous Weapon Systems and the Law of Armed Conflict*, 4 HARV. NAT'L SEC. J. 231 (2013); John O. McGinnis, *Accelerating AI*, 104 NW. U. L. REV. COLLOQUY 366 (2010).

10. Amanda Sharkey divides discussions on AWS into “arguments that focus on the extent to which AWS can adhere to IHL and the laws of war, and arguments that focus more on whether they should be used, even if they were shown to be capable of [adhering to laws].” Amanda Sharkey, *Autonomous Weapons Systems, Killer Robots and Human Dignity*, 21 ETHICS INFO. TECHN. 75, 76 (2019) (emphasis omitted). Some investigations of the first type also belong to the “evolution” side of the discussion—whether the use of AWS can adhere to the main principles of IHL as we know them. Cf., e.g., Yahli Shereshevsky, *International Humanitarian Law-Making and New Military Technologies*, 104 INT'L REV. RED CROSS 2131, 2143–45 (2022) (arguing that challenges posed by new technology can be best addressed by existing legal norms because law-making is not politically available); Rebecca Crotoft, *Regulating New Weapons Technology*, in THE IMPACT OF EMERGING TECHNOLOGIES ON THE LAW OF ARMED CONFLICT 3 (Ronald T.P. Alcalá & Eric Talbot Jensen eds., 2019) (“The aim of this chapter . . . is to step back and contemplate more generally whether and when new regulations are appropriate”); Kristen E. Eichensehr, *Cyberwar & International Law Step Zero*, 50 TEX. INT'L L. J. 355 (2015) (analyzing recent reactions to technological developments in warfare that occur more rapidly than international law); Peter Pascucci, *Distinction and Proportionality in Cyberwar: Virtual Problems with a Real Solution*, 26 MINN. J. INT'L L. 419 (2017) (suggesting an additional protocol to the Geneva Convention to address deficiencies regarding cyberwar in IHL); Jeffrey T. G. Kelsey, Note, *Hacking into International Humanitarian Law: The Principles of Distinction and Neutrality in the Age of Cyber Warfare*, 106 MICH. L. REV. 1427 (2007) (arguing that IHL should evolve through customary law to adjust to the challenges of cyber warfare); Dafna Dror-Shpoliansky & Yuval Shany, *It's the End of the (Offline) World as We Know It: From Human Rights to Digital Human Rights — A Proposed Typology*, 32 EUR. J. INT'L L. 1249 (2021) (exploring the need to reshape international human rights law to address the emergence of digital technologies).

11. See, e.g., U.S. MISSION INT'L ORGS. GENEVA, *U.S. Commentaries on the Guiding Principles* (Sept. 1, 2020), [hereinafter *U.S. Commentaries*] (Submission to the Group of Governmental Experts (GGE) on Lethal Autonomous Weapons Systems (LAWS)) (“International humanitarian law continues to apply fully to all weapons systems, including the potential development and use of lethal autonomous weapons systems.”).

others relevant actors insist on the need for developing new, specially tailored norms.¹²

In the same vein, much of the academic literature on AWS focuses on the practical ability of technological systems to comply with key IHL principles, such as distinction and proportionality; in particular, whether AWS are technologically capable of accurately distinguishing between combatants and civilians and avoiding incidental harm to civilians that is excessive relative to the military advantage anticipated.¹³ This is a crucial empirical question that underpins many dystopian scenarios of robots running amok and causing widespread havoc.

By contrast, this Article focuses on another set of normative investigations that is unique to the debate over the legality of AWS.¹⁴

12. See, e.g., U.N. Grp. of Gov't Experts, Proposal for a Mandate to Negotiate a Legally-Binding Instrument that Addresses the Legal, Humanitarian and Ethical Concerns posed by Emerging Technologies in the Area of Lethal Autonomous Weapons Systems (LAWS), Submitted by Austria, Brazil and Chile, U.N. Doc. CCW/GGE.2/2018/WP.7 (Aug. 30, 2018) (“[e]stablish[es] an open-ended Group of Governmental Experts to negotiate a legally-binding instrument to ensure meaningful human control over critical functions in lethal autonomous weapon systems.”).

13. See, e.g., Peter Asaro, *On Banning Autonomous Weapon Systems: Human Rights, Automation, and the Dehumanization of Lethal Decision-Making*, 94 INT'L REV. RED CROSS 687, 692 (2012) (“In terms of the legal acceptability of these systems under existing IHL, the primary question appears to be whether autonomous systems will be able to satisfy the principles of distinction and proportionality.”); Christof Heyns, *Autonomous Weapons in Armed Conflict and the Right to a Dignified Life: An African Perspective*, 33 S. AFR. J. ON HUM. RTS. 46, 52 (2017) (“There are several factors that could possibly impede the ability of autonomous weapons to operate according to the rule of distinction, including the technological inadequacy of existing sensors; a robot's inability to understand context; and the difficulty of translating IHL language and definitions of civilian and combatant into computer programming.”).

14. Another unique issue for AWS that is outside the scope of this essay is responsibility for violations of international law committed by autonomous weapons. See, e.g., Asaro, *supra* note 13, at 693 (describing concerns over attributing operator or state responsibility for AWS actions); See generally, Magdalena Pacholska, *Military Artificial Intelligence and the Principle of Distinction: A State Responsibility Perspective*, 53 ISR. L. REV. 3 (2023) (arguing that state responsibility applies to AWS actions); Russell Buchan & Nicholas Tsagourias, *Autonomous Cyber Weapons and Command Responsibility*, 96 INT'L L. STUD. 645 (2020) (proposing a form of commander liability for AWS actions); Crotoft, *supra* note 9 (suggesting the development of a legal regime of “war torts” to address AWS responsibility concerns); McFarland & McCormack, *supra* note 9 (discussing the responsibility of LAWS developers); Beard, *supra* note 9 (advocating a determination of accountability based on human judgement).

We consider the concept of *meaningful human control* as a new and potentially revolutionary legal requirement developed for regulating the use of AWS. The fact that discussions over the legality of AWS often center around this concept conveys a normative expectation that decisions potentially leading to the deprivation of human life—or to serious harm to other fundamental human rights—should be made by human beings or with the close involvement of human beings. This expectation suggests that the use of AWS should be prohibited in the absence of meaningful human control, even if they can be used in ways that are otherwise compatible with IHL or even more compatible with IHL than weapon systems operated by human soldiers.

Against this backdrop, Part I of this Article introduces and discusses the concept of meaningful human control and its normative implications. Following a brief review of existing approaches, we offer a novel moral argument against the use of fully autonomous weapon systems without meaningful human control. We maintain that humans' freedom to refrain from exercising the powers conferred on them by the laws of war is fundamental to human dignity and to maintaining humanity in warfare. The use of AWS may therefore run contrary to both the dignity principle and the fundamental humanitarian ethos underlying IHL. This discussion provides the normative underpinnings for legally requiring meaningful human control over lethal weapons systems.

Part II then examines whether a legal requirement of meaningful human control can be developed from existing legal frameworks, particularly IHL and IHRL. The second novel contribution of this Article is our claim that the requirement of meaningful human control may find some support in meta principles of IHL and even more support within IHRL. In particular, we consider the concept of human dignity in IHL and IHRL and focus on the exercise of choice, including opting to afford protections that go beyond those required by the strict letter of the law, as a component thereof. IHL is generally considered as the dominant law governing theatres of hostilities by virtue of the legal principle of *lex specialis*, which prioritizes the body of law more specifically tailored for a given situation. Nevertheless, we argue that IHRL might be recognized as the *lex specialis* for capturing novel humanitarian concerns, such as meaningful human control, that were not anticipated when IHL was codified.

To be sure, the question of whether AWS should be regulated by existing or new legal norms is more than an abstract exercise. Effective international governance of AWS is more likely, in the short run, if legal regulation can be derived from existing frameworks. The main legal policy forum for international negotiations on regulating the use of AWS is the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems (GGE LAWS), established in 2016 under the auspices of the Convention on Certain Conventional Weapons (CCW).¹⁵ Thus far, it appears unlikely that members of the GGE LAWS will agree to create a new treaty on AWS.¹⁶ In fact, they have found that even determining how to reinterpret existing norms requires significant political heavy lifting, and that consensus around meaningful human control has been elusive.¹⁷

Finally, Part III of the Article addresses the use of military AI beyond AWS. While initial discussions about military AI focused almost exclusively on AWS, the conversation is now expanding to address other AI systems, including those used for gathering intelligence and targeting decision support systems. Critics argue that automation bias and other factors may lead to undue deference to machine-generated recommendations in the application of these systems, undermining the meaningful nature of the human control that nominally attaches to them. We ask whether a meaningful

15. Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to be Excessively Injurious or to Have Indiscriminate Effects, Oct. 10, 1980, 1342 U.N.T.S. 137, 163.

16. It should be noted that the GGE LAWS is an ongoing process and there are continuous attempts by some states and other actors to promote a binding document on AWS. See Benjamin Perrin, *Lethal Autonomous Weapon Systems & International Law: Growing Momentum Towards a New International Treaty*, 29 AM. SOC. INT'L L. – INSIGHTS 1 (2025) (describing contemporary efforts in the GGE LAWS and beyond to create a treaty on AWS); GGE on LAWS (Nov. 8, 2024), [https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-_Group_of_Governmental_Experts_on_Lethal_Autonomous_Weapons_Systems_\(2024\)/Revised_rolling_text_as_of_8_November_2024_final.pdf](https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-_Group_of_Governmental_Experts_on_Lethal_Autonomous_Weapons_Systems_(2024)/Revised_rolling_text_as_of_8_November_2024_final.pdf) [https://perma.cc/6EEX-65LZ] (describing the current status of consensus agreement among members of the GGE LAWS).

17. See, e.g., Shereshevsky, *supra* note 10, at 2146 (discussing the challenges in the regulation of autonomous weapons); Charli Carpenter, *A Better Path to a Treaty Banning 'Killer Robots' Has Just Been Cleared*, WORLD POL. REV. (Jan. 7, 2022), <https://www.worldpoliticsreview.com/a-better-path-to-a-treaty-banning-ai-weapons-killer-robots/> [https://perma.cc/X3JU-YQDZ] (describing international failure to reach an agreement at the Convention on Conventional Weapons Review Conference).

human control requirement applies to these systems and argue that, at least from a human dignity perspective, there is an important normative distinction between AWS and AI-driven decision-supporting systems.

I. AUTONOMOUS WEAPONS AND MEANINGFUL HUMAN CONTROL – THE IMPORTANCE OF CHOICE

In this Part, we discuss the relationship between human agency, dignity, and the humanitarian ethos underlying IHL. We first introduce the concept of meaningful human control and the normative considerations supporting its central role in AWS. We then link meaningful human control to IHL's normative pragmatism and a variety of dignitary justifications. Finally, we propose that meaningful human control facilitates choice and agency, while comporting with IHL's basic structure. We believe that the position we present here ultimately raises the normative bar for relying on pragmatic claims to justify the use of AWS.

A. The debate over meaningful human control

The concept of meaningful human control has been the subject of much debate.¹⁸ It is a vague term that may be defined either broadly or narrowly. For example, human control could be interpreted broadly to include involvement of human beings in any stage of an AI system's life cycle, such as the design of the AWS or the decision to deploy it to a particular mission. In this article, we employ a narrower understanding of meaningful human control, involving *the ability of a human operator to intervene in the actual targeting decision of an autonomous weapon at some point during target engagement*—that is, a human in, or on, the operational targeting 'loop'.¹⁹ This working definition, which excludes human interactions

18. See, e.g., Linda Eggert, *Rethinking 'Meaningful Human Control,'* in RESPONSIBLE USE OF AI IN MILITARY SYSTEMS 213, 214–17 (Jan Maarten Schraagen ed., 1st ed. 2024) (challenging normatively the “widespread faith” in meaningful human control); Merel Ekelhof, *Moving Beyond Semantics on Autonomous Weapons: Meaningful Human Control in Operation*, 10 GLOB. POL'Y 343, 347 (2019) (“meaningful human control in relation to the human-machine relationship during the deployment of weapons is not the only, nor the most appropriate, approach to comprehensively address concerns of losing human control as it fails to take into account trivial factors of military practice”).

19. See Jean-Michel Verney, Thomas Vinçotte & Laurent le Quement, *Human-On-the-Loop*, in DELIVERING NATO AIR & SPACE POWER AT THE SPEED OF

before target engagement and after an attack takes place, allows us to focus on the core normative question about real-time human involvement in target identification and attack execution.

Some commentators urge that meaningful human control would have essential instrumental implications for AWS because it would enhance the proper functioning of these systems—preventing mistakes, malfunctions, and unpredictable outcomes.²⁰ Concerns about proper functioning and accuracy are reflected in much of the popular culture surrounding autonomous weapons, where, in many cases, the system fails to work as planned, leading to horrible consequences.²¹

Such terrible outcomes would constitute serious violations of IHL and IHRL norms. The requirement for meaningful human control is intended, from this perspective, to mitigate the risks emanating from potential technical failures or machine shortcomings and to increase compliance with international law.

A second body of literature relating to meaningful human control revolves around responsibility for violations of IHL, claiming that assigning moral and legal responsibility depends on human involvement in the life cycle of autonomous weapons.²² This normative position is reflected in the guiding principles of the GGE LAWS. The principle addressing responsibility explicitly refers to human control, stating that “[a]ccountability for developing, deploying and using any emerging weapons system in the framework

RELEVANCE: JOINT AIR & SPACE POWER CONFERENCE 131 (JAPCC ed. 2021) (discussing the human-on-the-loop terminology).

20. See, e.g., Abhimanyu George Jain, *Autonomous Weapon Systems, Errors and Breaches of International Humanitarian Law*, 21 J. INT’L CRIM. JUST. 1005 (2023); Rebecca Crootof, *A Meaningful Floor for “Meaningful Human Control,”* 30 TEMP. INT’L & COMP. L.J. 53 (2016) (discussing the need to clarify what level of human supervision is required for autonomous weapons).

21. See, e.g., *OUTSIDE THE WIRE*, (Automatik Entertainment et al. 2021) (A highly advanced AI-driven soldier (*Leo*) defects and determines that humans must be eliminated to prevent future wars.); *CHAPPIE* (Colombia Pictures et al. 2015) (A police robot gains sentience but is manipulated by criminals, leading to violent unintended consequences.); *I, ROBOT* (20th Century Fox Film Corp. et al. 2004) (The AI *VIKI* (Virtual Interactive Kinetic Intelligence) decides that the best way to protect humanity is to subjugate it, leading to a robot uprising.).

22. See, e.g., Asaro, *supra* note 13, at 693 (recognizing that in certain situations “no human individual could be held responsible for the actions of the autonomous weapon”); Matthew T. Miller, *Command Responsibility: A Model for Defining Meaningful Human Control Student Notes*, 11 J. NAT’L SEC. L. & POL’Y 533 (2020) (suggesting that commanders’ IHL duties can operationalize meaningful human control and ensure accountability for AWS use).

of the CCW must be ensured in accordance with applicable international law, including through the operation of such systems within a responsible chain of human command and control.”²³ The relationship between human control and individual responsibility is, however, located beyond the scope of this article, which focuses exclusively on non-instrumental justifications for requiring meaningful human control.

Assuming, as this article does, that AWS can comply with IHL while engaging in targeting operations as well as or better than humans, we focus on the importance of exercising meaningful human control *even when such control does not lead to a better compliance record*. Put differently, the question we address below pertains to the potential tension between pragmatic dimensions of warfare regulation (which aim at minimizing superfluous harm and unnecessary suffering) and the moral value of insisting on maintaining human control over targeting decisions, even if the involvement of human controllers does not reduce harm or suffering.

B. IHL and normative pragmatism

IHL is commonly understood as a legal regime that operates in a non-ideal world, in which the normal order of things has already collapsed due to the breakdown of norms on the prohibition of the use of force.²⁴ Under such extreme conditions, IHL is considered to be a fundamentally pragmatic legal discipline, built not on abstract or absolutist notions of justice, but rather on the practical need to regulate violent acts in ways that maximize the reduction of harm and suffering whenever possible. According to this understanding of IHL, the values underlying the legal regime do not necessarily lend support to a normative argument in favor of a meaningful human control requirement for AWS. In this section, we first address the legal regime’s normative pragmatism and explain why it might be viewed as counseling against such a requirement. We then move to other possible normative justifications.

23. U.N. Office for Disarmament Affairs, Meeting of the High Contracting Parties to the Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to Be Excessively Injurious or to Have Indiscriminate Effects: Final Report, U.N. Doc. CCW/MSP/2019/9, annex III, para. D (Dec. 13, 2019) [hereinafter GGE on LAWS Guiding Principles].

24. See, e.g., James Pattison, *The Case for the Nonideal Morality of War: Beyond Revisionism versus Traditionalism in Just War Theory*, 46 POL. THEORY 242, 242-63 (2018); John Rawls, *The Law of Peoples*, 20 U. CHI. PRESS 36 (1993).

One prominent example of IHL's normative pragmatism is its principle of belligerent equality—the idea that IHL applies equally to all parties in a conflict, regardless of the justice of their cause or the identity of the party responsible for initiating the hostilities.²⁵ Permitting asymmetry in warring parties' legal obligations likely would undermine compliance, hopelessly politicize the enforcement of IHL, and lead to increased harm and suffering. Therefore, this equal applicability formula sets aside questions of moral blameworthiness in order to reduce overall levels of harm and suffering. In this way, belligerent equality reflects a deeply pragmatic vision of how armed conflicts should be regulated, prioritizing the reduction of harm and suffering over the pursuit of absolute moral justice and ensuring that humanitarian protections continue to remain available even under the most extreme conditions.

In the same vein, scholars have argued that under certain conditions, AWS should be deemed legal or even legally required under IHL, even without meaningful human control. This may be the case if AWS can be placed within a framework of accountability and minimize human harm and suffering in armed conflict by complying with the main IHL norms just as well as, or even better than, humans.²⁶

25. See, e.g., Vaios Koutroulis, *And Yet It Exists: In Defence of the Equality of Belligerents Principle Hague International Tribunals: International Criminal Courts and Tribunals*, 26 LEIDEN J. INT'L L. 449, 449 (2013) (defining the principle of equality between belligerents as “the equal application of *jus in bello* to all belligerent parties of an international armed conflict irrespective of who violated *jus ad bellum*”); Adam Roberts, *The Equal Application of the Laws of War: A Principle Under Pressure*, 90 INT'L REV. RED CROSS 931, 932 (2008) (“[T]he laws of war apply equally to all belligerent parties in an international armed conflict[.]”); Jonathan Somer, *Jungle Justice: Passing Sentence on the Equality of Belligerents in Non-International Armed Conflict*, 89 INT'L REV. RED CROSS 655, 655 (2007) (defining equality of belligerents as including “the capacity of armed opposition groups to pass sentences on individuals for acts related to the hostilities”).

26. See, e.g., Eric Talbot Jensen, *Autonomy and Precautions in the Law of Armed Conflict*, 96 INT'L L. STUD. 577, 578 (2020) (arguing that autonomous systems may be capable of complying with IHL's functional requirements just as well as, or even better than, humans); Marco Sassoli, *Autonomous Weapons and International Humanitarian Law: Advantages, Open Technical Questions and Legal Issues to be Clarified*, 90 INT'L L. STUD. 308, 319–338 (2014) (discussing various tradeoffs of using autonomous systems instead of human judgement under IHL); Kenneth Anderson & Matthew Waxman, *Law and Ethics for Autonomous Weapon Systems: Why a Ban Won't Work and How the Laws of War Can*, in JEAN PERKINS TASK FORCE ON NATIONAL SECURITY & LAW ESSAY SERIES, STAN. U., HOOVER INST. 1, 11 (2013) (“[T]he fact of machine selection of target and

C. Dignitary justifications for meaning human control

Yet, some scholars have insisted on an independent requirement for meaningful human control. The literature that defends this requirement, independent of consequential considerations, often relies on the notion of human dignity. But the concept of human dignity is understood differently by different commentators. One set of scholars asserts that “death by algorithm” crosses a moral line and is against human dignity because AWS will be unable to understand or value the human lives that they were taking.”²⁷ This approach suggests that machines’ inability to acknowledge their victims’ humanity, and treatment of their human targets as mere objects, degrades their victims’ dignity.²⁸ Heller clarifies that this argument insists that only moral agents can make the decision to kill.²⁹

A second dignity-based set of claims maintains that human reflection and discretion are essential conditions for justifying decisions with lethal effects.³⁰ Asaro has articulated a legal position that insists on human decision-making for such reasons: “[H]uman judgement is constitutive of the system of justice. That is, if any system of justice is to apply to humans, then it must rely upon human reason. Justice itself cannot be delegated to automated processes.”³¹

One problem with both approaches is that they assume a clear distinction between lethal actions with and without meaningful

engagement with it—does not violate the law of armed conflict. Indeed, as the following sections discuss, it might turn out over time that for some purposes and forms of attack or defense, autonomous weapons may be able to be more discriminating and precise than human beings.”); Schmitt & Thurnher, *supra* note 9, at 234 (proposing that some autonomous systems, if properly developed and regulated, have the potential to better minimize harms and risks to civilians compared to non-autonomous systems).

27. Sharkey, *supra* note 10, at 83–84 (attributing this position to various scholars); Daniele Amoroso & Guglielmo Tamburrini, *Autonomous Weapons Systems and Meaningful Human Control: Ethical and Legal Issues*, 1 CURR. ROBOT. REP. 187, 189 (2020).

28. Sharkey, *supra* note 10, at 78.

29. Heller, *supra* note 9, at 6–11.

30. Sharkey, *supra* note 10, at 84. See Eliav Liebllich & Eyal Benvenisti, *The Obligation to Exercise Discretion in Warfare: Why Autonomous Weapons Systems Are Unlawful*, in AUTONOMOUS WEAPONS SYSTEMS: LAW, ETHICS, POLICY 245, 272 (Nehal Bhuta et al. eds. 2016) (positing similar arguments regarding human reflection and discretion based on notions from administrative law).

31. Asaro, *supra* note 13, at 701.

human control.³² But if the AWS programmers enable their technology to execute an attack that is legally sanctioned under IHL against a legitimate target, it is not clear what morally distinguishes the soldier's decision on the battlefield from that of the humans who programmed and deployed the AWS. The question is even more complex where the soldier on the frontline relies on information (e.g., attack coordinates) provided by a remote command and control center, which might be using decision supporting AI systems. Arguably, there must be a qualitative difference between targeting by AWS and targeting by soldiers in such scenarios that explains the moral intuition against AWS.

One qualitative difference may be “compassion-based.”³³ Some argue that the ability to feel—and especially the capacity for compassion and empathy—matters in war. They argue that, because AWS lack this capacity for compassion, they inflict harm and suffering in a morally deficient manner.³⁴ This idea finds support in the language on human compassion in General Comment 36 by the Human Rights Committee, which discusses the compatibility of AWS with the right to life.³⁵ As the Comment suggests, the ability to feel compassion toward civilians and enemy combatants is a key element in the humanization of war, which supplements the legal protections found in IHL. Arguably, replacing human operators with AWS would remove a layer of humanitarian protection that the law, supplemented by compassion principles, currently guarantees.

While compassion-based arguments have merit in our view, we note that those who base the case against the use of AWS on human emotional capacities must also contend with the counterargument that soldiers' emotions and mental inclinations, such as anger or cruelty, might lead to legal violations and excess

32. See Heller, *supra* note 9, at 6–19; Lieblich and Benvenisti, *supra* note 30, at 256–59.

33. Heller, *supra* note 9, at 57–58 (referring to this type of argument as arguments about compassion).

34. See, e.g., Markus Wagner, *The Dehumanization of International Humanitarian Law: Legal, Ethical, and Political Implications of Autonomous Weapon Systems*, 47 VAND. J. TRANSNAT'L L. 1371, 1414–16 (2014) (describing the debate of whether AWS can have compassion).

35. Human Rights Committee, General Comment No. 36, The Right to Life (Article 6), UN Doc. CCPR/C/GC/36 (Sept. 3, 2019), at para. 65 [hereinafter GC 36] (“For example, the development of autonomous weapon systems lacking in human compassion and judgment raises difficult legal and ethical questions concerning the right to life, including questions relating to legal responsibility for their use.”).

harm and suffering.³⁶ Still, as we explain below, compassion-based arguments capture something fundamental that does constitute a qualitative difference between targeting decisions with and without meaningful human control. They are therefore relevant to our own position regarding the moral significance of agency and choice in targeting operations.

Note that arguments about compassion often focus on compassion expressed towards enemy combatants whose lives are spared even though they constitute legitimate targets, rather than towards civilians who can never be lawfully targeted (but may be treated, in certain circumstances, as collateral harm). Examples of this dynamic appear in the famous “naked soldiers” section of Michael Walzer’s *Just and Unjust Wars*, where Walzer describes scenarios in which soldiers chose not to shoot enemy combatants despite having the legal right to do so.³⁷ For example, he quotes from George Orwell’s memoir of the Spanish Civil War, in which Orwell describes an incident he experienced while serving in an anarchist battalion:

At this moment a man, presumably carrying a message to an officer, jumped out of the trench and ran along the top of the parapet in full view. He was half-dressed and was holding up his trousers with both hands as he ran. I refrained from shooting at him. It is true that I am a poor shot and unlikely to hit a running man at a hundred yards . . . Still, I did not shoot partly because of that detail about the trousers. I had come here to shoot at “Fascists;” but a man who is holding up his trousers isn’t a “Fascist,” he is visibly a fellow-creature, similar to yourself, and you don’t feel like shooting at him.³⁸

Walzer takes Orwell’s story and other examples to suggest that they illustrate a vital insight about the regulation of warfare:

It is not against the rules of war as we currently understand them to kill soldiers who look funny, who are taking a bath, holding up their pants, reveling in the sun, smoking a cigarette. The refusal of these five

36. See Heller, *supra* note 9, at 59 (“Moreover, insofar as the possibility of compassion is offered as an argument against potentially more discriminating autonomous weapons, the calculus must take into account all of the situations in which a human soldier fails to exercise compassion and intentionally kills a civilian.”).

37. MICHAEL WALZER, *JUST AND UNJUST WARS: A MORAL ARGUMENT WITH HISTORICAL ILLUSTRATIONS* 138–43 (4th ed. 2006).

38. *Id.* at 140.

men, nevertheless, seems to go to the heart of the war convention. For what does it mean to say that someone has a right to life? To say that is to recognize a fellow creature . . . whose person is as valuable as my own.³⁹

This is not purely a compassion-centered argument. Walzer does not claim that a capacity for emotion is crucial for minimizing harm and suffering in warfare (thus avoiding the need to address the counterargument regarding negative human emotions, such as anger or fear). Rather, he highlights how the socio-legal relations constituted by war remain just one facet of human interaction on the battlefield. This approach underscores the ability of soldiers to step outside the strict logic dictated by the war paradigm to preserve their humanity and personal agency by deciding not to kill enemy soldiers.

D. Meaningful human control and the importance of choice

Walzer's "naked soldier" stories vividly illustrate our core argument: The unique capacity for choice distinguishes human targeting decisions from algorithmic ones, over and beyond questions of specific moral or emotional motivations. In the examples provided by Walzer, both the attacking and the attacked individuals are soldiers; yet, the real possibility that either soldier might deviate from the script attached to the roles they play in the military conflict suggests that war, despite its brutality, is not a closed system of predetermined outcomes. Indeed, the mere presence of human soldiers as decision-makers—notwithstanding the significant organizational and sociological pressures they are subject to—allows them to say, "I will not pull the trigger," for reasons of compassion, morality or another impulse which may be unforeseen, unexpected, and unexplained. Such features of open-endedness in decision-making disappear when an AWS executes lethal force without meaningful human control. Rather than leaving room for last-minute deviations or inexplicable changes of course and heart, AWS renders outcomes fixed, fueling a sense of determinism for all of those involved.

In exploring why the loss of such choice over targeting decisions is so troubling, Alex Leveringhaus emphasizes that "Killer Robots *qua* artificial agents lack a central component of human

39. *Id.* at 142. Walzer is building on an observation by Jean-Jacques Rousseau. JEAN-JACQUES ROUSSEAU, *DU CONTRAT SOCIAL, OU PRINCIPES DU DROIT POLITIQUE* 13–15 (1762) (describing Rousseau's observation that war is not between man and man, but between the State and the State).

agency, the ability to do otherwise: *not* to shoot the target.”⁴⁰ He then explains why the capacity to deviate from a predetermined course of action matters:

[T]he enemy about to be targeted is still a fellow human being with one life to live. Killer Robots, by contrast, do not relate to humans in the way humans relate to each other. They do not have the agency to put their gun down because ‘they pity the poor sod’ they have been programmed to target. Notwithstanding centuries of human carnage, I think retaining human agency at the point of force delivery, thereby protecting the freedom not to pull the trigger, push the button, or throw a grenade, is essential for retaining our humanity in exactly the situation that challenges it the most: war.⁴¹

Christof Heyns strikes a similar chord, underscoring how AWS deprive the battlefield of any “[i]t’s lawful, but it’s awful” moments.⁴² In one passage of an article that was partly based on his work as a UN Special Rapporteur on Extrajudicial, Summary or Arbitrary Executions, Heyns writes:

Where it is legally or even ethically permissible to use force, humans may decide not [sic] do so because something holds them back. This possibility is excluded when autonomous weapons are used. . . . If autonomous weapons can be calibrated to use force to the full extent that the law allows humans[,] . . . lives that could have been spared will be lost because . . . the potentially restraining elements of humanity [are removed].⁴³

We share these concerns but emphasize that our argument against AWS centers on the disappearance of choice itself. Once AWS entirely replaces the attacking human soldier, neither the human deploying the system nor the human facing its lethal consequences can reasonably believe that a different outcome remains possible. The battlefield thus becomes an arena where a predetermined chain of events plays out, divorced from any vestige of human agency or hope that lives will be spared in a moment of compassion or even due to sheer whim. By contrast, when humans remain in control of lethal

40. Alex Leveringhaus, *What’s So Bad About Killer Robots?*, 35 J. APPLIED PHIL. 341, 349 (2018).

41. *Id.* at 350 (footnote omitted).

42. Heyns, *supra* note 13, at 63.

43. *Id.* at 63 (footnote omitted).

weapon systems, a residue of open-ended possibility endures, preventing warfare from fully descending into a strictly preordained and pre-scripted process of organized lethal violence.

E. Agency as a component of dignity

The debate over the use of AWS involves fundamental aspects of the relationship between warfare and humanity. People's control over their lives and life stories is a crucial aspect of what it means to be human.⁴⁴ Indeed, the question of free will has been a major theological puzzle in many religions and a central philosophical quandary debated over the ages.⁴⁵ In the midst of battle, when humanitarian harm and suffering are pervasive and extensive, we are of the opinion that retaining some degree of autonomous choice is vital for soldiers in order to preserve their humanity and the human features of the situation in which they find themselves.

Moreover, retaining the possibility for an open-ended future is not merely a feature of humanity; it is intimately tied to human dignity as a core moral value of human society. Feeling that one's life circumstances—particularly in matters of life and death—are never entirely predetermined underscores a basic human need for meaning. Indeed, a key element of human dignity is the conviction that one can, at least in principle, autonomously shape one's life story. Joseph Raz, for example, underscores the significance of personal autonomy by articulating that “[t]he ruling idea behind the ideal of personal autonomy is that people should make their own lives. The autonomous person is a (part) author of his own life. . . . [It] is the vision of people controlling, to some degree, their own destiny,

44. See generally KENNON M. SHELDON, *FREELY DETERMINED: WHAT THE NEW PSYCHOLOGY OF THE SELF TEACHES US ABOUT HOW TO LIVE* (2022) (discussing the importance of free will as part of the human experience); A. Will Crescioni et al., *Subjective Correlates and Consequences of Belief in Free Will*, 29 *PHILOS. PSYCHOL.* 41 (2016) (presenting empirical evidence that free will can be associated with meaningfulness of life); ROBERT KANE, *THE SIGNIFICANCE OF FREE WILL* (1996) (exploring the various philosophical and theoretical underpinnings of free will). But see DERK PEREBOOM, *FREE WILL, AGENCY, AND MEANING IN LIFE* (2014) (presenting a skeptical view of free will and its relationship to morality).

45. See generally Timothy O'Connor & Christopher Franklin, *Free Will*, in *THE STANFORD ENCYCLOPEDIA OF PHILOSOPHY* (Edward N. Zalta & Uri Nodelman eds., 2022), <https://plato.stanford.edu/archives/win2022/entries/freewill/> [<https://perma.cc/V9QL-JZ9P>] (providing a broad overview of the philosophical landscape surrounding the idea of free will).

fashioning it through successive decisions throughout their lives.”⁴⁶ This open-endedness assures us that we are not mere cogs in a machine, and that our choices, however small or spontaneous, still matter.

When we imagine a dystopian battlefield fully controlled by AWS, a principal source of concern is precisely the elimination of any “way out” of the warfare script, leaving no room for a last-minute reprieve. Instead, the outcome of every military engagement appears fixed in advance, fueling an acute sense of determinism and fatalism and depriving participants of the sense that life is shaped by human choices that entail moral implications and responsibility. In doing so, it strikes at the existential bedrock of what makes events in human life—including involvement in human warfare—more than a mere mechanical unfolding of predetermined consequences.

The “naked soldiers” stories highlight these human characteristics from the targeting soldier’s perspective. The key aspect of these stories is, in our view, not merely about possessing a capacity for an emotional response but preserving a sense of agency that transcends the logic of war and the laws of war. From the viewpoint of those individuals lawfully targeted by lethal force, agency and choice give rise to the hope or belief that their fate is not sealed.

Significantly, an important element of a dignified life is not just the *actual* ability to exercise agency; it also includes the *perception* that because all parties to human interactions possess agency, no one’s life story is pre-scripted.⁴⁷ This perception goes to the heart of the qualitative difference between AWS and human soldiers or weapon systems employed under meaningful human control. It explains, we believe, much of the moral intuition against AWS.

46. JOSEPH RAZ, *THE MORALITY OF FREEDOM* 369 (1986). For discussions of personal autonomy as an aspect of human dignity in the legal context *see, e.g.*, Jacob Weinrib, *Human Dignity and Autonomy*, in *MAX PLANCK ENCYCLOPEDIA OF COMPARATIVE CONSTITUTIONAL LAW* (Rainer Grote, Frauke Lachenmann & Rüdiger Wolfrum eds., 2020); Frederick Schauer, *Speaking of Dignity*, in *THE CONSTITUTION OF RIGHTS: HUMAN DIGNITY AND AMERICAN VALUES* 178 (Michael J. Meyer & William A. Parent eds., 1992). *But see* Jennifer Nedelsky, *Laws Relations: A Relational Theory of Self, Autonomy, and Law* (2011).

47. *See* SAUL SMILANSKY, *FREE WILL AND ILLUSION* (2000) (arguing that belief in free will is essential, regardless of whether free will actually exists).

F. Meaningful human control and the structure of IHL

A second significant aspect of the capacity to choose whether to exercise legal powers to apply lethal force involves the laws of war themselves. We suggest that the basic structure of IHL contrasts with the way AWS operate. Therefore, even if such systems could comply with each specific IHL norm, the general ethos of IHL counsels against their use. While Leveringhaus discusses agency and choice in moral terms, we propose that they are also central to the laws regulating armed conflicts. The central aspect of agency we identify here involves the exercise of choice. Such a choice is central for the *jus in bello* and perhaps also for broader rules on the use of force—*jus ad bellum*—and IHRL. Significantly, the laws of war do not *obligate* soldiers to use force against legitimate targets; rather, they *permit* doing so under certain conditions. The ultimate choice whether to use force remains in the hands of the soldier, and others in the chain of command, who retains agency to choose not to engage the target.

This empowering, but not mandating, framework reflects the fundamental ethos of IHL, wherein resorting to violence may be factually necessary and thus legally permissible, but always a compromise of sorts. This compromise is a by-product of the non-ideal reality in which armed conflicts take place. For this reason, the laws of war never demand, nor encourage, attacking every legitimate target during an armed conflict; rather, they leave warring parties free to choose whether, and to what extent, they will utilize their right to attack legitimate targets.⁴⁸ This discretionary structure helps preserve the notion of war as an exceptional state of human affairs and allows for moderating considerations of humanity in the most dire circumstances.⁴⁹ It also reduces the prospects that blind application of IHL would lead to a total or almost-total war, even one that is only conducted towards legitimate targets.⁵⁰ By contrast,

48. See ANNE QUINTIN, THE NATURE OF INTERNATIONAL HUMANITARIAN LAW: A PERMISSIVE OR RESTRICTIVE REGIME? 2–10 (2020).

49. Cf. *Corfu Channel (U.K. v. Alb.)*, Judgment, 1949 I.C.J. 4, 22 (Apr. 9) [hereinafter *Corfu Channel*] (alluding to “elementary considerations of humanity” as a general principle applicable in times of war and peace).

50. Military AI beyond AWS, such as target identification rules, were recognized as a potential target identification multiplier. If there is nothing that prevents them from targeting all legitimate targets, there is a real danger of significant escalation of the pain and suffering of armed conflicts. See, e.g., Omar Yousef Shehabi & Asaf Lubin, *Israel – Hamas 2024 Symposium – Algorithms of War: Military AI and the War in Gaza*, Articles of War (Jan. 24, 2024), <https://lieber.westpoint.edu/algorithms-war-military-ai-war-gaza/>

deploying AWS generates predetermined outcomes that could transform the battlefield from one featuring the selective use of permissible force into one that potentially brings to bear the full extent of permissible force, all the time, against all legitimate targets.

While the framework of analysis offered here provides a principled argument in favor of requiring meaningful human control over targeting decisions, we should underscore two important caveats to our framework. First, we do not maintain that other considerations could never override our argument in favor of meaningful human control—particularly if AWS can dramatically out-perform human operators and considerably reduce harm and suffering in and around the battlefield. Still, the position we present here significantly raises, we believe, the normative bar for relying on pragmatic claims to justify the use of AWS.⁵¹

Second, it might be possible to create AWS that can fully adhere to IHL while mimicking human behavior and refraining from exercising force in circumstances where the strict letter of the law permits it. If such a technological development were to become a new reality, the strength of the argument we present here would diminish. The development might weaken our argument grounded in the permissive aspects of IHL.⁵² Still, there are few indications of concrete plans by states to develop such technology.⁵³

[<https://perma.cc/B74G-23NU>]; Noah Sylvia, *The Israel Defense Forces' Use of AI in Gaza: A Case of Misplaced Purpose*, RUSI (July 4, 2024), <https://www.rusi.org/explore-our-research/publications/commentary/israel-defense-forces-use-ai-gaza-case-misplaced-purpose> [<https://perma.cc/52AV-K826>].

51. See generally Larry Alexander, *Deontology at the Threshold Festschrift in Honor of Michael Moore's Placing Blame*, 37 SAN DIEGO L. REV. 893 (2000) (offering a seminal account of threshold deontology and law); Leveringhaus, *supra* note 40, at 348–53 (applying threshold deontology to the case of AWS).

52. The impact of the development on our agency and choice argument would be less clear, especially since any acts over and beyond the strict letter of the law would themselves be pre-scripted and governed by algorithmic programming.

53. Even if future advances in AI technology would make such a development technically feasible, it remains unclear whether states will have an incentive to move in such direction and whether it will be possible to preserve compliance with IHL rules and prohibitions while avoiding certain undesirable features of mimicking human behavior (which may include not only the capacity for the exercise of restraint but also harmful tendencies to use excessive force due to fear, revenge or hatred). Note also that mimicking human behavior could raise difficult questions about the moral agency of ‘socialized’ AWS. See DAVID J. GUNKEL, *THE MACHINE QUESTION: CRITICAL PERSPECTIVES ON AI, ROBOTS, AND ETHICS* (2012).

With these two caveats in mind, the next section explores whether we can find a doctrinal basis for a meaningful human control requirement within the existing IHL and IHRL regimes, given the normative argument for the requirement we have presented.

II. IHL, IHRL, LEX SPECIALIS AND MEANINGFUL HUMAN CONTROL

In this Part, we examine whether a legal requirement of meaningful human control can be developed from existing legal frameworks, particularly IHL and IHRL. We claim that a requirement of meaningful human control could find some support in meta principles of IHL and even more support within IHRL. We consider, in this connection, the concept of human dignity as it exists in IHL and IHRL and the extent to which it incorporates the exercise of choice, including the choice to afford protections that go beyond those required by the strict letter of the law. We also discuss in this Part the normative relationship between IHL and IHRL and claim that, while IHL is generally considered as the *lex specialis* in battlefield situations, IHRL might be recognized as the *lex specialis* for capturing novel humanitarian concerns, such as meaningful human control, that were not anticipated when IHL was codified.

A. A doctrinal basis for a meaningful human control requirement under IHL

If a meaningful human control requirement can be grounded in existing international law, it would most likely be located in either IHL or IHRL—the two bodies of law providing for the protection of human lives in times of armed conflict. As noted above, since AWS are a new form of military technology used for operational targeting, IHL is the natural starting point for evaluating the legality of means or methods of war. However, identifying IHL rules that explicitly mandate meaningful human control proves challenging. As noted above, we assume in this article that core IHL principles, such as those related to distinction, proportionality, and precautions, can be respected without resort to direct human control or oversight.⁵⁴ Indeed, it cannot be ruled out that AWS might develop a capacity to distinguish between combatants and civilians or even undertake

54. See generally Eric Talbot Jensen, *The (Erroneous) Requirement for Human Judgement (and Error) in the Law of Armed Conflict*, 96 INT'L L. STUD. 26 (2020) (arguing that IHL does not require direct human control over weapons systems and documenting states' divergent views on the issue).

proportionality analysis (according to the best practices in the field) with greater accuracy than humans. In the absence of a specific agreement to outlaw AWS, it is difficult to assert that, under existing law, its use is prohibited under *lex lata*.⁵⁵

The principle of humanity, as a general principle of the laws of war, might support a claim that IHL should be read to require meaningful human control over AWS. The interpretative move needed here involves maintaining that, in the absence of a clear rule of *lex lata*, either allowing or prohibiting AWS—i.e., facing a legal lacuna—law-interpreters can resort to the principle of humanity by invoking the Martens clause.⁵⁶ This clause, which is more than one hundred years old, refers to “the laws of humanity, and the dictates of the public conscience” as possible gap-fillers.⁵⁷ The possibility of using the Martens clause as part of the humanization of IHL has received some support in international law scholarship, although the precise scope of the clause and its normative implications remain contested.⁵⁸ Alternatively, it is possible to rely on “elementary considerations of humanity,” i.e., on the principle of humanity itself, as a gap-filling general principle of law.⁵⁹

Such interpretations of IHL rely on our previous normative discussion, which links agency and choice to the principle of humanity and suggests that meaningful human control is integral to maintaining humanity in warfare. Further, our point regarding the importance of viewing IHL as a set of permissions to use force in certain cases—rather than a set of obligations or encouragements to use force—can also be seen as aligned with the principle of humanity.

55. For example, the ICRC’s position on AWS emphasizes ethical considerations, notes legal considerations, and calls on states to agree to ban some of their uses. See *ICRC Position on Autonomous Weapon Systems*, INT’L COMM. OF THE RED CROSS (MAY 12, 2021), <https://www.icrc.org/en/document/icrc-position-autonomous-weapon-systems> [https://perma.cc/477X-DKJ4]

56. Convention No. IV, Respecting the Laws and Customs of War on Land, The Hague, Oct. 18, 1907, pmbl., 205 Consol. T.S. 277; Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of International Armed Conflicts (Protocol I), art. I(2), June 8, 1977, 1125 U.N.T.S. 3 [hereinafter AP I].

57. *Id.*

58. See, e.g., Theodor Meron, *The Martens Clause, Principles of Humanity, and Dictates of Public Conscience*, 94 AM. J. INT’L L. 78 (2000) (describing the clause’s centrality in progressing humanization of international humanitarian law); Antonio Cassese, *The Martens Clause: Half a Loaf or Simply Pie in the Sky?*, 11 EUR. J. INT’L L. 187 (2000) (describing the ambiguous implications of the Martens clause).

59. Corfu Channel, *supra* note 49, at 22.

This approach suggests that the choice whether to resort to legally permitted force operates as a form of humanitarian restraining factor in a context where maintaining humanity is extremely challenging.

While it is commonly assumed that IHL principles do not by themselves create new obligations for states—especially in areas of IHL typically regulated by explicit treaty provisions⁶⁰—it is not completely unprecedented for law-interpreters to rely on general principles of international law to establish concrete legal obligations. The ICRC Interpretive Guidance on the Notion of Direct Participation in Hostilities serves as a primary example of such an attempt.⁶¹ The authors of the Interpretive Guidance faced a doctrinal challenge similar to the one before us—there is no explicit IHL rule that requires combatants to capture, rather than kill, persons directly participating in hostilities.⁶² Therefore, the authors applied principles of humanity and military necessity to establish within IHL an obligation to use the least harmful means, even against legitimate targets.⁶³

B. Legal risks associated with relying on general principles in IHL

Using general principles to create new and concrete obligations for states that go beyond pre-existing *lex lata* carries with it, however, two primary legal risks. First, compliance with new norms based on general principles is expected to be rather limited if they are rejected by key states and commentators comprising the IHL legal community (as had largely happened to the ICRC Interpretive Guidance).⁶⁴ The debate over whether AWS can outperform humans,

60. See Legality of the Threat or Use of Nuclear Weapons, Advisory Opinion, 1996 I.C.J. 226, 248 (July 8) (“Nor, however, is there any principle or rule of international law which would make the legality of the threat or use of nuclear weapons or of any other weapons dependent on a specific authorization.”)

61. ICRC, INTERPRETIVE GUIDANCE ON THE NOTION OF DIRECT PARTICIPATION IN HOSTILITIES UNDER INTERNATIONAL HUMANITARIAN LAW (Nils Melzer ed., 2009) [hereinafter INTERPRETIVE GUIDANCE].

62. Several of the aforementioned naked soldiers stories revolve around situations where an enemy soldier, who can be lawfully targeted under IHL throughout the duration of the armed conflict, does not represent an immediate threat to the adversary and can easily be captured instead of being killed (e.g., a naked soldier swimming in a river). See WALZER, *supra* note 37, at 138–43.

63. INTERPRETIVE GUIDANCE, *supra* note 61, Principle IX.

64. See, e.g., W. Hays Parks, *Part IX of the ICRC Direct Participation in Hostilities Study: No Mandate, No Expertise, and Legally Incorrect Forum: Direct Participation In Hostilities: Perspectives on the ICRC Interpretive Guidance*, 42

and as a result reduce overall harm and suffering in and around the battlefield, renders it unlikely that a consensus would emerge over the application of the principle of humanity to AWS. Still, using general principles for requiring meaningful human control differs significantly from the ICRC Interpretive Guidance. While it is commonly understood that there was no *lex lata* obligation under IHL to use the least harmful means before the Interpretative Guidance was published,⁶⁵ the characterization of IHL as a legal regime based on a permission to use force rather than on an obligation to do so is uncontested. Our claim is that this meta-IHL principle is an important component of the principle of humanity and that it may entail requiring meaningful human control over AWS. In addition, the GGE LAWS issued, by way of consensus, eleven guiding principles regarding the regulation of LAWS, which contain a vague allusion to meaningful human control in paragraphs (c) and (d).⁶⁶ While these guiding principles are ambiguous in language and non-binding in nature, they nonetheless suggest state support for some requirement of meaningful human control, including explicit support from some states.⁶⁷

N.Y.U. J. INT'L L. & POL. 769, 783–85 (2009) (offering doctrinal criticism of the ICRC's Interpretive Guidance).

65. See, e.g., Ryan Goodman, *The Power to Kill or Capture Enemy Combatants*, 24 EUR. J. INT'L L. 819 (2013) (suggesting that IHL contains an implicit norm of harm minimization requiring fights to wound or capture, rather than kill, enemy combatants); Michael N. Schmitt, *Wound, Capture, or Kill: A Reply to Ryan Goodman's "The Power to Kill or Capture Enemy Combatants,"* 24 EUR. J. INT'L L. 855 (2013) (counterarguing that IHL creates no express or implied obligation to wound or capture, rather than kill, enemy combatants).

66. Guiding Principles affirmed by the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons System, para. (c), (d), U.N. Doc. CCW/MSP/2019/9 (2019), (“(c) Human-machine interaction, which may take various forms and be implemented at various stages of the life cycle of a weapon, should ensure that the potential use of weapons systems based on emerging technologies in the area of lethal autonomous weapons systems is in compliance with applicable international law, in particular IHL. In determining the quality and extent of human-machine interaction, a range of factors should be considered including the operational context, and the characteristics and capabilities of the weapons system as a whole; (d) Accountability for developing, deploying and using any emerging weapons system in the framework of the CCW must be ensured in accordance with applicable international law, including through the operation of such systems within a responsible chain of human command and control”).

67. For example, the joint commentary of Austria, Belgium, Brazil, Chile, Ireland, Germany, Luxembourg, Mexico, and New Zealand underscores the importance of human-machine interaction for ensuring compliance with IHL and accountability. See Joint ‘Commentary’ on Guiding Principles A, B, C and D,

The second, more serious risk in establishing specific legal requirements based on general principles of IHL is that doing so may create more uncertainty around the laws of war. To be sure, clear rules of conduct increase compliance in most circumstances, including military operations undertaken in the most challenging environments of adversity and uncertainty, characterized by the proverbial “fog of war.”⁶⁸ Furthermore, the interpretive community of IHL includes both humanitarian actors, who seek to expand IHL’s protective role, and military lawyers, who often emphasize the opposite requirements of military necessity. The opposing interests of these interpreters may mean that opening the door for broad teleological interpretations of IHL based on general principles will legitimize a broad reading of both the principle of humanity and the principle of military necessity. Reliance on the latter principle might be used to justify legal measures not explicitly prohibited under IHL (e.g., targeting financial assets of a militant group or withholding consent to relief operations), which would allow parties to a conflict to inflict more harm and suffering than what *lex lata*, as we currently understand it, permits. The influence of states on the interpretation of IHL makes this concern particularly significant. Still, our argument may limit the risk of an expansive use of general principles because it draws on the basic structure of IHL, whereas other attempts to use general principles impute those principles from outside of the IHL.

In light of these considerations, we believe that creating specially tailored IHL rules to regulate AWS and to mandate meaningful human control is a preferable regulatory path to reinterpretation of IHL.⁶⁹ As we have noted, such acts of law-creation

submitted by Austria, Belgium, Brazil, Chile, Ireland, Germany, Luxembourg, Mexico, and New Zealand (Sep. 2020) <https://documents.unoda.org/wp-content/uploads/2020/09/GGE20200901-Austria-Belgium-Brazil-Chile-Ireland-Germany-Luxembourg-Mexico-and-New-Zealand.pdf> [<https://perma.cc/E87B-HRWQ>]. But see U.S. COMMENTARIES, *supra* note 11, at 3 (describing that the U.S. position includes the possibility of deploying fully autonomous weapon systems without express intention of a human operator).

68. See Amichai Cohen, *Rules and Standards in the Application of International Humanitarian Law*, 41 ISR. L. REV. 41, 62–65 (2008) (emphasizing that clear, pre-specified rules are crucial in IHL to limit post hoc discretion and ensure more predictable compliance).

69. Indeed, the UN Secretary General and the President of the ICRC called on states to formulate a specific agreement restricting the use of LAWS. See *Joint Call by the United Nations Secretary-General and the International Committee of the Red Cross to Establish Prohibitions and Restrictions on Autonomous Weapon Systems*, INT’L COMM. RED CROSS (May 10, 2023), <https://www.icrc.org/>

have not happened to date, and the prospects of them ever happening remain unclear. Given the high moral stakes of using AWS and the rapid advancement of AWS technologies, waiting for future legal developments in the GGE or other lawmaking body poses significant risks. Against this background, we now proceed to explore whether IHRL could offer a more appropriate legal framework for establishing an obligation for meaningful human control over AWS.

C. A doctrinal basis for a meaningful human control requirement under IHRL

Since human dignity is a core moral value from which many human rights derive,⁷⁰ it is unsurprising that dignity-based objections to the use of AWS involve, sooner or later, exploring the legality of AWS under IHRL.⁷¹ While some aspects of human dignity are also protected by IHL (e.g., prohibition of humiliation⁷² and unnecessary suffering⁷³), others are not (e.g., the aforementioned absence of an explicit capture rule). By contrast, dominant human rights theories regard human rights as giving expression to human dignity and associated values such as human agency, autonomy, liberty, solidarity, and capabilities.⁷⁴ If the use of AWS violates human dignity by transforming human interactions involving life and death, agency, choice and hope, into mechanical interactions that preclude the ability to act beyond the strict letter of the law, then IHRL may have something to say on the legality of AWS.

en/document/joint-call-un-and-icrc-establish-prohibitions-and-restrictions-autonomous-weapons-systems [https://perma.cc/5PBA-V4AG].

70. See G.A. Res. 217 (III) A, Universal Declaration of Human Rights, at 71 (Dec. 10, 1948) [hereinafter UDHR].

71. See, e.g., Asaro, *supra* note 13 at 693 (questioning the sufficiency of the IHRL to regulate AWS).

72. See, e.g., AP I, *supra* note 56, at art. 75(2)(b) (“The following acts are and shall remain prohibited at any time and in any place whatsoever, whether committed by civilian or by military agents: . . . outrages upon personal dignity, in particular humiliating and degrading treatment, enforced prostitution and any form of indecent assault”).

73. See, e.g., *id.* at art. 35(2) (“It is prohibited to employ weapons, projectiles and material and methods of warfare of a nature to cause superfluous injury or unnecessary suffering”).

74. See generally, Christopher McCrudden, *Human Dignity and Judicial Interpretation of Human Rights*, 19 EUR. J. INT’L L. 655 (2008) (explaining the various jurisdiction-specific applications of human dignity); Oscar Schachter, *Human Dignity as a Normative Concept*, 77 AM. J. INT’L L. 848 (1983) (exploring the meaning of human dignity).

However, doctrinal sources of analysis pointing to the dignity-based objections to AWS under IHRL remain somewhat limited. One pathbreaking scholar who has started going in this direction is Christof Heyns, who has proposed that meaningful human control could be grounded in the right to life, which he understood as including a right to a *dignified* life.⁷⁵ Heyns contends that the right to life necessitates deliberative human decision-making when using weapon systems.⁷⁶ He further justified a meaningful human control requirement by relying on the notion of a dignified life, suggesting that someone subject to the force of an AWS “is reduced to being an object that has to be destroyed. . . . They have no avenue, futile or not, of appealing to the humanity of the enemy, or hoping their humanity will play a role because it is a machine on the other side.”⁷⁷

Heyns’s broad reading of the right to life as including a right to a dignified life is, *inter alia*, based on the African Commission on Human and Peoples’ Rights General Comment 3 on the Right to Life (which he helped draft).⁷⁸ This reading received further support in 2019, when the Human Rights Committee (of which Heyns and one of the authors of this Article were members at the time) accepted his position and stated in General Comment 36 on the Right to Life (GC 36) that “[t]he right to life is a right that should not be interpreted narrowly. It concerns the entitlement of individuals to be free from acts and omissions that are intended or may be expected to cause their unnatural or premature death, as well as to enjoy a life with dignity.”⁷⁹ The drafters of GC 36 also recognized “the central importance to human dignity of personal autonomy”⁸⁰ and expressed doubts concerning the compatibility of AWS with the right to life with dignity. Indeed, GC 36 provides the following:

States parties engaged in the deployment, use, sale or purchase of existing weapons and in the study, development, acquisition or adoption of weapons, and means or methods of warfare, must always consider their impact on the right to life. For example, the

75. Christof Heyns, *Human Rights and the Use of Autonomous Weapons Systems (AWS) during Domestic Law Enforcement*, 38 HUM. RTS. Q. 350, 378 (2016).

76. *Id.* at 373.

77. Heyns, *supra* note 13, at 63.

78. African Commission on Human and Peoples’ Rights, *General Comment 3 on the African Charter on Human and Peoples’ Rights: The Right to Life* (Article 4) (Nov. 18, 2015).

79. GC 36, *supra* note 35, at para. 3.

80. *Id.* at para. 9.

development of autonomous weapon systems lacking in human compassion and judgment raises difficult legal and ethical questions concerning the right to life, including questions relating to legal responsibility for their use. The Committee is therefore of the view that such weapon systems should not be developed and put into operation, either in times of war or in times of peace, unless it has been established that their use conforms with article 6 and other relevant norms of international law.⁸¹

The General Comment focuses on problems of compassion, judgment, and legal responsibility—all featured in the existing scholarship on meaningful human control. While our novel emphasis on agency and choice is not explicitly mentioned in the General Comment, its list of “difficult legal and ethical questions” to which it alludes is clearly not an exhaustive one.⁸² Arguably, insisting on the ability to refrain from acting as authorized by the law aligns with the articulation of the relationship between personal autonomy, human dignity, and judgment in different parts of GC 36.

D. The relationship between IHL and IHRL in the Context of AWS

While some authors have argued for reliance on IHRL as the doctrinal basis for a meaningful human control requirement, one crucial element, sometimes overlooked in this discussion, is the legal relationship between IHL and IHRL in the context of AWS. While both legal frameworks are relevant to the use of AWS, it appears that IHRL offers a more solid doctrinal basis for limiting their application without meaningful human control. This immediately raises questions about their manner of normative interaction—that is, how to go about if IHL permits what IHRL prohibits or *vice versa*.

The co-application of IHL and IHRL is well-studied and exceeds the scope of this article.⁸³ It is sufficient to note that in situations where both legal regimes apply—such as armed conflicts involving measures falling under the jurisdiction of the participating

81. *Id.* at para. 65.

82. *Id.*

83. See, e.g., Cordula Droege, *The Interplay Between International Humanitarian Law and International Human Rights Law in Situations of Armed Conflict*, 40 ISR. L. REV. 310, 311 (2007) (explaining the abundance of overlap between the two disciplines).

states⁸⁴—the principle of systemic integration invites interpretive efforts to reconcile all relevant norms.⁸⁵ If reconciliation proves impossible, normative priority should be given, according to the ICJ, to the *lex specialis*—which, in situations concerning the conduct of hostilities, is typically IHL.⁸⁶ Still, the farther from the actual battlefield the situation at hand is, the greater the relevance of IHRL as a normative framework. In practice, all the elements that underlie co-application debates are often hotly contested, including whether an armed conflict exists, whether states have jurisdiction over individuals affected by the measures they take, whether IHL and IHRL can be reconciled, and what should be the dominant legal framework in particular cases. This is especially the case in murky legal situations, such as those involving low-intensity armed conflicts, situations of belligerent occupation, counter-terrorism operations, and cyber-attacks.⁸⁷ In recent years, IHRL has been considered the *lex specialis* in contexts which are significantly addressed under IHRL but which receive minimal attention under IHL, including privacy and data protection.⁸⁸ In these domains, scholars claim that IHL's silence creates an opening for the application of IHRL, even during armed conflicts and in operational contexts.⁸⁹

When applying this paradigm to the use of AWS, IHL appears to dominate the legal analysis we see in practice. The paradigm's IHL

84. See Yuval Shany, *The Extraterritorial Application of International Human Rights Law*, in 409 COLLECTED COURSES OF THE HAGUE ACADEMY OF INTERNATIONAL LAW / RECUEIL DES COURS 21-23 (2020) (analyzing the scope of territorial and extraterritorial application of IHRL to armed conflicts).

85. Vienna Convention on the Law of Treaties (VCLT), art. 31(3)(c), May 23, 1969, 1155 U.N.T.S. 331, (entered into force Jan. 27, 1980). See Marko Milanovic, *A Norm Conflict Perspective on the Relationship between International Humanitarian Law and Human Rights Law*, 14 J. CONFLICT & SEC. L. 459 (2009) (offering an approach focused on systemic integration).

86. Legality of the Threat, *supra* note 60, at para. 25.

87. See, e.g., David Kretzmer, *Rethinking the Application of IHL in Non-International Armed Conflicts*, 42 ISR. L. REV. 8, 41-42 (2009) (evaluating the application of IHL and IHRL to certain types of non-international armed conflicts).

88. Asaf Lubin, *The Rights to Privacy and Data Protection Under International Humanitarian Law and Human Rights Law*, in RESEARCH HANDBOOK ON HUMAN RIGHTS AND HUMANITARIAN LAW: FURTHER REFLECTIONS AND PERSPECTIVES 462 (Robert Kolb, Gloria Gaggioli & Pavle Kilibarda eds., 2022).

89. Tal Mimran & Yuval Shany, *Integrating Privacy Concerns in the Development and Introduction of New Military or Dual-Use Technologies*, in THE RIGHTS TO PRIVACY AND DATA PROTECTION IN TIMES OF ARMED CONFLICT 44 (Russell Buchan & Asaf Lubin eds., 2022).

norms governing weapon development and the conduct of hostilities apply in particular. Indeed, the focused attention from civil society groups, such as Article 36, on weapon reviews pursuant to Article 36 of the First Additional Protocol to the Geneva Convention (AP1), as well as the focus of inter-state bodies such as the GGE LAWS on the regulatory role of IHL rather than IHRL, highlights the perceived centrality of IHL in the legal regulation of AWS.⁹⁰

However, while IHRL has almost never been applied to the core conduct of hostilities such as targeting decisions, we suggest that there is a strong argument that IHRL gives rise to a meaningful human control requirement for AWS. The silence of IHL with respect to meaningful human control (which can be described as involving neither clear support nor opposition to the concept), in combination with IHRL's support for this requirement, suggests that no real conflict exists between the two legal regimes in this context. Put differently, we suggest that IHL does not deal with the dignitary aspects of AWS (in the same way that it does not deal with privacy and data protection) and that IHRL can be therefore seen as the *lex specialis* in this specific context.

While this argument may be legally meritorious, its prospects for acceptance by states remain uncertain. We note in this regard that the existence of an IHRL-based obligation for meaningful human control is itself in doubt, since even GC 36 identified difficult questions but did not provide definitive answers. GC 36 recommended a human rights impact assessment—comparable to the weapon review undertaken pursuant to article 36 of the First Additional Protocol—without predetermining the outcome of the process.⁹¹

Finally, like our normative argument in Part I of this article, we note that most IHRL norms, including the right to life (the legal basis upon which the Human Rights Committee identified a possible requirement for meaningful human control) are relative in nature.⁹² This means that even if there are strong reasons to insist on a meaningful human control requirement, there may be overriding reasons pointing in the other direction—including AWS' superior

90. *Autonomous weapon systems: Evaluating the capacity for 'meaningful human control' in weapon review processes*, ARTICLE36: DISCUSSION PAPER (Nov. 2017), <https://article36.org/wp-content/uploads/2013/06/Evaluating-human-control-1.pdf> [<https://perma.cc/UR7C-KW8Q>].

91. GC 36, *supra* note 35, at para. 65.

92. See UDHR, *supra* note 70, at art. 29(2) (explaining that the exercise of rights and freedoms are subject to limitations determined by law and other societal norms like morality, public order, and general welfare).

adherence to IHL than what is expected from a human weapons operator. While this may put a heavy burden on states to justify the use of AWS, it may also further complicate initiatives to push for a sweeping ban of these weapon systems.

III. MEANINGFUL HUMAN CONTROL AND MILITARY AI BEYOND AWS

The discussion around military AI has recently expanded beyond AWS. The integration of military AI tools into targeting processes—excluding their actual engagement with a specific target—has drawn significant scholarly and popular attention.⁹³ While full-fledged AWS largely remain a future phenomenon, military AI beyond AWS play a role in contemporary conflicts. These tools, which support targeting decisions in different ways and degrees—such as through collection and analysis of intelligence on potential targets and issuing targeting recommendations—are already being used by or integrated into several militaries at an accelerating pace. They raise significant legal and ethical concerns.⁹⁴

93. See, e.g., Anthony King, *Digital Targeting: Artificial Intelligence, Data, and Military Intelligence*, 9 J. GLOB. SEC. STUD. 2, 3 (2024) (discussing AI's role in enhancing military intelligence analysis and targeting processes); Merel A. C. Ekelhof, *AI is Changing the Battlefield, but Perhaps Not How You Think: An Analysis of the Operationalization of Targeting Law and the Increasing Use of AI in Military Operations*, in RESEARCH HANDBOOK ON WARFARE AND ARTIFICIAL INTELLIGENCE 161 (Robin Geiß & Henning Lahmann eds., 2024); Jon R. Lindsay, *War is from Mars, AI is from Venus: Rediscovering the Institutional Context of Military Automation*, 7 TEX. NAT'L SEC. REV. 29, 31 (2024) (explaining how improvements in AI lead to discussions on "accurate targeting, unintended civilian casualties, and meaningful human control"); H. W. Meerveld et al., *The Irresponsibility of Not Using AI in the Military*, 25 ETHICS INFO. & TECH. 14, 1–6 (2023) (discussing the ethical implications of AI use in military operations and the associated risks of military AI in the broader context); FORREST E. MORGAN ET AL., MILITARY APPLICATIONS OF ARTIFICIAL INTELLIGENCE: ETHICAL CONCERNS IN AN UNCERTAIN WORLD 107–112 (RAND Corp. ed., 2020) (discussing the public acceptance of military AI in offensive, defensive, and escalating military operations).

94. Most of the accounts about the use of military AI focus on the conflict in Gaza post Oct. 7, 2023, but there are also accounts of such uses in other conflicts, such as the war in Ukraine. See, e.g., Harry Davies & Yuval Abraham, *Revealed: Israeli Military Creating ChatGPT-Like Tool Using Vast Collection of Palestinian Surveillance Data*, THE GUARDIAN (Mar. 6, 2025), <https://www.theguardian.com/world/2025/mar/06/israel-military-ai-surveillance> [<https://perma.cc/B3LC-TWXD>] (explaining Israel's use of a powerful AI tool for surveillance and spying in the Israel-Hamas war in Gaza); Scott Nover, *In Ukraine's AI-enabled War Against Russia, Humans Still Call the Shots*, GZERO MEDIA (Mar. 11, 2025), <https://www.gzeromedia.com/gzero-ai/in-ukraines-ai->

The conflict in Gaza often serves as a key example of the use of AI-enabled Decision Support Systems (DSS). Concerns about the large number of civilian casualties in Gaza have led many commentators to suggest that DSS are a significant part of the problem.⁹⁵ We cannot replicate here the entire debate over DSS and their potential contribution to increased civilian casualties. It suffices to state that we believe it is important to carefully examine the respective contribution to destructive outcomes during war of multiple causes, including the various implications of using DSS. In the Gaza context, it is possible that other factors—such as anger, hate, and vengeance, loose legal interpretations, and the entrenchment of militants within the civilian population during a high intensity urban conflict—has contributed to the very high number of civilian casualties. It is also possible that factors unique to DSS played a significant role. For example, civilian casualties might be heightened due to DSS’ ability to rapidly identify a large number of targets, the limited accuracy of these systems, or their contribution

enabled-war-against-russia-humans-still-call-the-shots [https://perma.cc/TL83-YVCV] (discussing how Ukraine has advanced in software development and procured AI-enabled drones during the Russian-Ukraine Conflict); Samuel Bendett & David Kirichenko, *Battlefield Drones and the Accelerating Autonomous Arms Race in Ukraine*, MODERN WAR INSTITUTE (Jan. 10, 2025), <https://mwi.westpoint.edu/battlefield-drones-and-the-accelerating-autonomous-arms-race-in-ukraine/> [https://perma.cc/END7-5DLK] (describing Ukraine’s battlefield as a “war lab for the future” because of its testing of new autonomous weaponry); *Israel Defence Forces’ Response to Claims About use of ‘Lavender’ AI Database in Gaza*, THE GUARDIAN (Apr. 3, 2024), <https://www.theguardian.com/world/2024/apr/03/israel-defence-forces-response-to-claims-about-use-of-lavender-ai-database-in-gaza> [https://perma.cc/5PKV-XS6S] (explaining that the IDF uses AI systems for target identification in the Israeli-Palestine war); ISR. DEF. FORCES, *The IDF’s Use of Data Technologies in Intelligence Processing* (June 18, 2024), <https://www.idf.il/210062> [https://perma.cc/2D7T-925J] (explaining that IDF uses AI tools for the intelligence distillation process to identify targets).

95. See, e.g., Emelie Andersin, *The Use of the ‘Lavender’ in Gaza and the Law of Targeting: AI Decision Support Systems and the Use of Facial Recognition Technology*, J. INT’L HUMANITARIAN LEGAL STUD. 1, 13–14 (2025) (recognizing that the speed at which DSS produces recommendations may induce military commanders to act more quickly, which prevents them from making necessary assessments to reduce casualties); Luke Moffett & Jessica Dorsey, *The Warification of International Humanitarian Law and the Artifice of Artificial Intelligence in Decision-Support-Systems: Restoring Balance Through the Legitimacy of Military Operations*, 2024 YEARBOOK OF INT’L HUMANITARIAN L. (forthcoming, 2025) (explaining how the “lack of understandability” of AI-DSS decision-making creates operational challenges that may lead to unexpected or excessive civilian harm).

to operators' potential moral distancing. In this article, however, we refer only to the applicability of our principled argument against the absence of meaningful human control over AWS to military AI beyond AWS.

The broader discussion around military AI mirrors the debate surrounding AWS in many respects. Much of the debate addresses whether such tools can facilitate compliance with IHL given the complexity of IHL's requirements and the technological challenges of programming decision-support systems to apply these rules properly.⁹⁶ As with discussions on AWS, this article does not address these intricate compliance questions that pose serious concerns whenever military AI is considered. Instead, it focuses on the parallels and distinctions between debates over AWS and debates about meaningful human control in the context of AI-based decision support systems.

Scholars writing on the topic have questioned whether human operators can truly exercise meaningful control over targeting decisions that rely on input generated by AI systems. Although humans must nominally approve or abort any targeting operation, concerns about automation bias and the speed of decision-making on the basis of data produced by rapidly operating AI systems suggest that human participation may be insufficient to ensure effective oversight and control.⁹⁷ For example, imagine an intelligence-gathering tool that tells a commander the precise location of a weapons cache in a residential building and states that it is certain

96. See, e.g., Neil C. Renic & Elke Schwarz, *Inhuman-in-the-loop: AI-Targeting and the Erosion of Moral Restraint*, OPINIOJURIS (Dec. 19, 2023), <https://opiniojuris.org/2023/12/19/inhuman-in-the-loop-ai-targeting-and-the-erosion-of-moral-restraint/> [https://perma.cc/J79H-F2J9]; Robert Stewart & Georgia Hinds, *Algorithms of War: The use of Artificial Intelligence in Decision Making in Armed Conflict*, HUMANITARIAN LAW AND POLICY (Oct. 24, 2023), <https://blogs.icrc.org/law-and-policy/2023/10/24/algorithms-of-war-use-of-artificial-intelligence-decision-making-armed-conflict/> [https://perma.cc/5EHF-5MSZ]; Tal Mimran & Gal Dahan, *Artificial Intelligence in the Battlefield: A Perspective from Israel*, OPINIOJURIS (Apr. 20, 2024), <https://opiniojuris.org/2024/04/20/artificial-intelligence-in-the-battlefield-a-perspective-from-israel/> [https://perma.cc/58SW-V3Z6]; ANNA NADIBAIDZE, INGILD BODE & QIAOCHU ZHANG, AI IN MILITARY DECISION SUPPORT SYSTEMS: A REVIEW OF DEVELOPMENTS AND DEBATES (2024).

97. See, e.g., Marta Bo & Jessica Dorsey, *Symposium on Military AI and the Law of Armed Conflict: The 'Need' for Speed - The Cost of Unregulated AI Decision-Support Systems to Civilians*, OPINIOJURIS (Apr. 4, 2024), <https://opiniojuris.org/2024/04/04/symposium-on-military-ai-and-the-law-of-armed-conflict-the-need-for-speed-the-cost-of-unregulated-ai-decision-support-systems-to-civilians/> [https://perma.cc/T6MY-SXYT].

that no civilians are currently present there. How likely is the commander to question this recommendation amid high-intensity armed conflict? Such scenarios reflect the concerns appearing in much of the recent scholarship on these military AI systems, particularly regarding their use in ongoing hostilities.⁹⁸

However, the extent of deference given to such AI tools is an open empirical question that has not been thoroughly studied. While automation bias and other conflict-related factors may lead users to defer to AI, algorithmic aversion might prompt hesitation in relying on AI for life-or-death decisions. One study by Horowitz and Kahn investigated the interplay between automation bias and algorithmic aversion among members of the general public tasked with surveillance identification.⁹⁹ Their findings were mixed: participants tended to trust highly trained humans over highly trained AI but placed greater confidence in less-trained AI than in less-trained humans.¹⁰⁰ They also discovered that a higher level of familiarity with AI increased the likelihood of relying on it. In another study, Whyte explored AI's effect on elite decision-making in cyber conflicts, finding that users were more willing to rely on AI when it simply assisted human intelligence-gathering, as opposed to situations in which AI-generated information was the sole basis for an operational decision.¹⁰¹ Horowitz and Lin Greenberg similarly showed that national security experts were less likely to initiate military action when AI performed the relevant intelligence analysis, compared to human analysts.¹⁰² We note in this regard that besides the question of over-reliance on AI, there is also a risk that military commanders will over-rely on intelligence officers, and such tendencies should be factored into any normative evaluation.¹⁰³

98. See *id.*; Klaudia Klonowska, *Israel-Hamas 2024 Symposium – AI-Based Targeting in Gaza: Surveying Expert Responses and Refining the Debate*, LIEBER INST. (June 7, 2024), <https://lieber.westpoint.edu/ai-based-targeting-gaza-surveying-expert-responses-refining-debate/> [<https://perma.cc/U6K6-NGY3>].

99. Michael C. Horowitz & Lauren Kahn, *Bending the Automation Bias Curve: A Study of Human and AI-Based Decision Making in National Security Contexts*, 68 INT'L STUD. Q. 1 (2024).

100. *Id.*

101. Christopher Whyte, *Learning to Trust Skynet: Interfacing with Artificial Intelligence in the Cyberspace*, 44 CONTEMP. SEC. POL'Y 308 (2023).

102. Michael C. Horowitz & Erik Lin-Greenberg, *Algorithms and Influence: Artificial Intelligence and Crisis Decision-Making*, 66 INT'L STUD. Q. 1, 5 (2022).

103. See generally Asaf Lubin, *The Reasonable Intelligence Agency*, 47 YALE J. INT'L L. 119 (2022) (arguing that IHL's regulation of military intelligence is essential because military leaders tend to shift blame onto intelligence).

This article proceeds on the (questionable) assumption that concerns about automation bias are well-founded, and that human operators may indeed defer significantly to AI-based decision-support tools in targeting processes. If that is the case, and if these AI tools prove at least as capable as human operators in applying core IHL rules, the question remains whether they should be subject to the same normative and doctrinal framework that applies to AWS.

We propose that two significant differences may distinguish these two military AI cases, both of which go to the core of the agency and choice considerations discussed in Part I of this article. First, and most importantly, it is useful to examine the nature of deference toward military AI. It seems reasonable to assume that any deference arises from a belief that AI-based decision-support tools surpass humans in accurately determining whether a potential target is legitimate under IHL. However, the qualitative distinction between humans and AWS identified earlier rests on the ability to act over and beyond the strict letter of the law—to choose not to “pull the trigger” even when such course of action is permitted, based on considerations unrelated to legality or concerns about intelligence accuracy. It does not need to involve a strong moral intuition or positive emotion such as compassion. What matters, in our view, is the mere ability to choose, reflecting human agency. If that is indeed the key difference between AWS and human-operated weapon systems, a choice to defer to military AI would not necessarily run contrary to notions of human agency. Consequently, the normative claims that we advance above do not directly translate to the uses of military AI beyond AWS, if such uses are based on the ability to exercise choice whether to follow machine recommendations.

Second, our approach focuses on the formal ability to choose, rather than on the extent to which the choice is free from external pressures and influence. Note that even setting military AI aside, soldiers in armed conflict situations face substantial social and psychological pressures—including, at times, peer pressure to violate the law—that could influence their decisions.¹⁰⁴ In the chaotic

104. See Saira Mohamed, *Abuse by Authority: The Hidden Harm of Illegal Orders*, 107 IOWA L. REV. 2183 (2022) (reconceptualizing illegal orders as abuse of authority affecting a subordinate’s agency); MATTHEW TALBERT & JESSICA WOLFENDALE, *WAR CRIMES: CAUSES, EXCUSES, AND BLAME* (2019) (describing the situational pressures, including military culture and ideology, under which military personnel act); Ziv Bohrer, *Is the Prosecution of War Crimes Just and Effective? Rethinking the Lessons from Sociology and Psychology*, 33 MICH. J. INT’L L. 749, 763 (2012) (analyzing how cognitive and social constraints shape

conditions of warfare, one might question whether soldiers' on-the-spot decisions truly reflect independent deliberation or an effective exercise of control. Yet, we typically consider these soldiers to have enough agency to bear criminal responsibility for their actions.¹⁰⁵ Criminal liability for actions taken in one's capacity as a soldier underscore the value society places on agency and choice, even in situations where external pressures—social, psychological, and hierarchical—significantly constrain one's decision-making freedom.

The same principles apply to the possibility of automation bias in military AI. While soldiers may feel compelled to defer to an AI recommendation, they do not lose the formal capacity to second-guess or override that recommendation if they choose to. Similarly, individuals targeted by these soldiers do not lose all hope or belief that their lives might be spared. That agency disappears when humans have no role at all in the final targeting decision, as with AWS.

Our discussion of military AI beyond AWS is, as noted, confined to the central normative issue explored in this article. We do not reject the substantial concerns associated with using AI support systems for targeting—particularly, whether such systems produce sufficiently accurate intelligence and whether humans can adequately verify the information they generate. These are empirical and ethical questions that lie beyond this article's scope; indeed, the legality of such systems mostly depends on affirmative answers regarding their accuracy and reliability. Our narrower claim presented here is simply this: Unlike AWS—where an independent normative argument exists against the use of such systems even if they match or exceed human performance under IHL—other military AI tools remain subject to human direction and control at a level that allows for some degree of human agency and choice.

perceived moral agency in war); Martha Minow, *Living Up to Rules: Holding Soldiers Responsible for Abusive Conduct and the Dilemma of the Superior Orders Defence*, 52 MCGILL L. J. 1 (2007) (pointing out the role of the military in mitigating this effect).

105. Vasiliki Zoi Festa, *Erasing the 'Victim' from the 'Victim-Perpetrator': Expressivist Messages in the Ongwen Trial Judgment*, 00 J. INT'L CRIM. JUST. 1, 13 (2024) (considering the victim-perpetrator dichotomy of a "child soldier turned warlord"); *Prosecutor v. Mrkšić & Šljivančanin*, IT-95-13/1-A, Appeal Judgment, para. 94 (May 5, 2009); Rome Statute of the International Criminal Court, art. 33, July 17, 1998, 37 I.L.M. 999; Paola Gaeta, *The Defence of Superior Orders: The Statute of the International Criminal Court versus Customary International Law*, 10 EUR. J. INT'L L. 172 (1999) (comparing the conditional and absolute liability approaches to the "obedience to orders" defense).

CONCLUSION

This paper makes three main contributions to the extensive literature on AWS. First, the principle of human dignity includes the ability to refrain from exercising legal powers conferred on soldiers. This is because the choice to say “no” is an important aspect of human agency and choice. By contrast, AWS without meaningful human control set off pre-determined courses of events without any hope for those targeted. A meaningful human control requirement aligns with the ethos of IHL, which seeks to prohibit certain conduct while sanctioning force in other cases, but never imposes obligations to use lawful lethal force.

Second, we suggest that IHL and IHRL can each support a requirement for meaningful human control grounded in humanity and dignity, respectively. We argue that there are strong reasons to prefer IHRL and provide a new perspective on the relationship between these two bodies of law, proposing that IHRL may serve as the *lex specialis* in certain circumstances, even in relation to core conduct-of-hostilities issues.

Finally, we offer a distinction between cases with no agency and control whatsoever, and other cases, such as those involving decision-supporting military AI tools distinct from AWS, indicating that the rationale for a requirement of meaningful human control does not equally apply to situations involving and not involving AWS.

The non-ideal nature of warfare regulation perpetuates a well-known tension between military necessity and humanitarian considerations. This tension carries the risk that the compromises inherent in IHL may legitimize and normalize lawful harm and suffering in war. Walzer’s concept of the “naked soldier” reminds us of the essential humanity of soldiers, even though they are deemed legitimate targets. Being soldiers is not their only defining characteristic; rather, they remain human beings. Our article, beyond offering three novel contributions to the debate over autonomous weapon systems (AWS), seeks to underscore this by emphasizing IHL as a legal regime that permits, rather than compels, the use of force. Such an understanding has an expressive value well beyond our current discussion, and it is worthwhile to explore its potential normative implications in relation to other areas of IHL and IHRL.

The policy implication of our approach is that resort to AWS can only be justified, if at all, after a strong case is made in favor of the likelihood that its use would improve existing levels of compliance

with IHL to an extent that would compensate for the negative humanity and dignity implications of substituting human with machines as the final link in the targeting decision-making chain. The natural forum for making such a case would be weapon review processes under article 36 of API, or ones modeled thereafter. Alternatively, one can envision the gradual acceptance by states of some version of the ‘meaningful human control’ standard as a necessary safeguard against the broad negative consequences of use of full-fledged AWS we allude to in this article.