Going Beyond the "Killer Robots" Debate

Six Dilemmas Autonomous Weapon Systems Raise

Anzhelika Solovyeva, Nik Hynek

The debate on and around "killer robots" has been firmly established at the crossroads of ethical, legal, political, strategic, and scientific discourses. Flourishing at the two opposite poles, with a few contributors caught in the middle, the polemic still falls short of a detailed, balanced, and systematic analysis. It is for these reasons that we focus on the nitty-gritties, multiple pros and cons, and implications of autonomous weapon systems (AWS) for the prospects of the international order. Moreover, a nuanced discussion needs to feature the considerations of their technological continuity vs. novelty. The analysis begins with properly delimiting the AWS category as fully autonomous (lethal) weapon systems, capable of operating without human control or supervision, including in dynamic and unstructured environments, and capable of engaging in independent (lethal) decision-making, targeting, and firing, including in an offensive manner. As its primary goal, the article aims to move the existing debate to the level of a first-order structure and offers its comprehensive operationalisation. We propose an original framework based on a thorough analysis of six specific dilemmas, and detailing the pro/con argument for each of those: (1) (un)predictability of AWS performance; (2) dehumanization of lethal decision-making; (3) depersonalisation of enemy (non-)combatant; (4) human-machine nexus in coordinated operations; (5) stra-

Anzhelika Solovyeva, Nik Hynek. Going Beyond the "Killer Robots" Debate: Six Dilemmas Autonomous Weapon Systems Raise. *Central European Journal of International and Security Studies* 12, no. 3: 166–208.

tegic considerations; (6) AWS operation in law(less) zone. What follows are concluding remarks.

Keywords: autonomous weapon systems, killer robots, lethal decision-making, military ethics, artificial intelligence, security regulation, humanitarian law, revolution in military affairs, military strategy

Anzhelika Solovyeva
Nik Hynek

The speculative term "killer robots" has increasingly been penetrating into ethical, legal, political, strategic, scientific and academic discourses. "Robots" in this collocation are designated as a "colloquial rendering for autonomous weapon systems" (AWS).[1] These are in turn delimited by the International Committee of the Red Cross as weapon systems that "can independently select and attack targets, i.e. with autonomy in the 'critical functions' of acquiring, tracking, selecting and attacking targets,"[2] and by the United States Department of Defense as weapon systems "that, once activated, can select and engage targets without further intervention by a human operator."[3]

Discussions in relation to the integration of this technology into combat "have ranged from the moral and legal implications, to technical and operational concerns, to issues about international security and worries about cyber vulnerability."[4] A truly "multi-dimensional" debate[5] arose primarily around the central ethical concern of the moral and legal acceptability of delegating "to a machine or automated process the authority or capability to initiate the use of lethal force independently of human determinations of its moral and legal legitimacy"[6] – as accurately emphasised as the primary object of the debate, for example, by Peter Asaro and as will be elaborated in more detail below in this paper. This debate has been largely flourishing at the two distinct points and involving multiple actors. Currently, there are more than sixty non-governmental organisations (NGOs) in the international campaign calling for an international legally binding treaty to prohibit the development and use of AWS.[7] States are also involved into the discourse and the main venue for United Nations deliberations on AWS is the Convention on Certain Conventional Weapons in Geneva.[8] Military personnel, scientists and lawyers, ethicists and philosophers have contributed to the discussion.[9] While critics call for a blanket preventive global ban on the development, production and use of this warfare technology,[10] some proponents insist that, on the contrary, there may exist a "moral imperative" for their deployment in

combat,[11] and that a blanket prohibition may bring serious humanitarian risks, considering the possibility that AWS may potentially become more discriminating and more ethically preferable to alternatives.[12]

It is worth claryfying that these apparently incompatible positions do not exist independently of each other. Largely, the pieces of literature produced by the representatives of the two positions make it clear that both sides in the debate do recognise the existence and are aware of the content of couter-arguments to their own visions of the problematic. They both indeed use a set of axioms, raise certain hypotheses and issues to impugn the contradictory point of view. So no great monologue: it is the proper debate. This article neither aims to take sides in the debate or to question any of its building blocs intellectually, nor does it aim to draw a line between the two camps agent-wise or to simply retell the debate. Rather, the goal is to move "beyond" the existing discussion, in the following ways:

a. This analysis aims to move the debate to the level of a *first-order structure* by suggesting the authors' vision of a comprehensive operationalisation of the debate. Based on the writings of both proponents and opponents of AWS development and deployment as well as on other related sources and analytical articles, the paper aims to reveal multiple argument – counter-argument chains in relation to certain matters of the dispute in a multi-issue manner and a parallel counter-posing manner. The latter, within the context of this paper, means not counter-posing certain agents' visions against one another, but substantially counter-posing arguments themselves that in some way appeared in a related discussion by any of the cited authors. The focus on details in this analysis will help to demonstrate the full-fledged, double-sided nature of the key aspects of the debate.
b. For achieving the just-mentioned, this analysis aims to go beyond the broader categories of legality, morality, ethicality, military utility, political and strategic implications of "killer robots," which the debate is largely being built around. Instead it constructs the framework of *six specific dilemmas*, potentially raised by autonomous weapon systems. Each consists of constituent parts or supporting matters of the dispute. By analytically separating the dilemmas and (sub)sections, this paper aims to highlight that each of them represents a certain controversial

aspect of the debate. However, at the same time these multiple aspects are considered and presented in this paper as interlinked and intertwinned to form one common picture with regard to the nature of AWS and the challenges they pose. It is important to note that the way these dilemmas and their components are specified and structured in this paper, with each aspect potentially deserving its own place in future research, will also provide the reader with an analytical framework that may make it easier to detect and analyse multiple interlinkages between these aspects in greater depth and in multiple possible ways, which is beyond the scope of this article due to space limitation.

Delimiting the object of the debate

The robotic revolution in military affairs is currently underway as the next paradigm shift in the nature of warfare following the introduction of gunpowder and nuclear bombs.[13] Robot is a powered machine that senses, thinks and acts.[14] Robotic systems can (be) operate(d) semi- or fully-autonomously but they "cannot depend entirely on human control."[15] These are currently widely present in the modern battlefield and the functions vary from providing intelligence gathering, surveillance, reconnaissance, target acquisition and designation to engagement capabilities.[16]

To specify though, "robots are not weapons systems until they are armed"[17] and specifically a certain category of "armed robots"[18] or "robot weapons"[19] form the core of this debate. Robot weapons are not new to the modern combat. Most of the currently deployed armed robots today are semi-autonomous (e.g., the United States Air Force's Predator), but some autonomous systems are also emerging (e.g., the United States Navy's Phalanx Close-In Weapon System).[20] However, fully autonomous robot weapons "do not yet exist."[21] The debate is built around the idea of a preventive ban. It is, thus, fundamental for the purpose of this article to clarify the boundaries between the so-called "killer robots" and other armed robots that have been used in the battlefield.

Fully autonomous (lethal) weapon systems

We begin by bringing together a variety of collocations used to refer to the debated technological category, going beyond the emotive, one-sided, and analytically unproductive term "killer robots". Our ac-

ademic preference lies with a more neutral alternative term "autonomous weapon systems." Other options include lethal autonomous weapon systems,[22] lethal autonomous systems,[23] autonomous lethal technologies,[24] lethal autonomous weapons,[25] lethal autonomous robot weapons,[26] fully autonomous weapons,[27] fully autonomous armed robots,[28] fully autonomous robotic weapons[29] and others. They all allow for a capture of the two most distinctive features of this specific robot weapon sub-category, which encompasses "*fully*" and "*lethal*" autonomous weapon systems.[30] These two features differentiate the debated category from the existing armed robotic systems.

While *remotely operated systems*, primarily including "drones and unmanned ground and underwater vehicles," feature "systems based on robotic technologies," can be used offensively[31] and can be "lethally armed,"[32] they may be more correctly described as "uninhabited" rather than unmanned systems, although they are referred to in either way.[33] This is because their autonomous mission is primarily "to navigate, but not select and engage targets, autonomously"[34] and they only "enable those who control lethal force not to be physically present when it is deployed."[35] In turn, AWS will "add a new dimension to this distancing" where in addition to being physically removed from the kinetic action, humans will also become more detached from decisions to fire/kill and their execution.[36] Importantly, such systems will "eliminate human judgement in the initiation of lethal force."[37] AWS will close the gap between uninhabited and unmanned warfare.[38]

It is worth noting that some weapon systems are already "able to identify, track and engage incoming targets on their own" and "can already be set up so that humans are cut out of decision-making."[39] However, they represent only "the precursors" to the capabilities that will appear in future autonomous systems.[40] *Defensive weapon systems* are currently the only type of autonomous robots that have been deployed and they can only fire on targets within well-delimited areas, therefore, they can be seen as "extensions of electric fences."[41] Besides being solely defensive weapons, which are stationary or fixed, and are designed to operate within tightly set parameters and time frames, such systems are primarily pre-programmed to fire at inanimate targets.[42] Counter-rocket, anti-missile and anti-aircraft systems represent this definition in practice.[43] Although there also exists SGR-A1 "robotic stationary platform designed to replace or to assist South Korean sentinels in the surveillance of the demilitarized zone between North and

South Korea" and it can operate in an unsurpervised "mode" whereby, importantly, also "any human being detected there is classified as a target," this system is also only a precursor to the debated robot weapon sub-category because, although potentially expected to initiate autonomous lethal force, it is designed to operate in a strictly structured environment, i.e. "Korean demilitarized zone" where human access is "categorically prohibited."[44] Importantly, with regard to the systems described in this paragraph, humans still "decide when and where to deploy the weapon, and can intervene to prevent its operation."[45] In turn, AWS will "operate without human control or supervision in dynamic, unstructured, open environments, attacking various sets of targets, including inhabited vehicles, structures or even individuals," potentially being able to "learn and adapt their behavior."[46]

To summarise, "killer robots" or, better still, true autonomous weapon systems (AWS), in their proper meaning and as referred to in the context of this article, can be differentiated from all other robot weapon categories by a unique combination of features defining their category: (1) they are *fully autonomous*,[47] including (a) their ability to *operate without human control or supervision in dynamic, unstructured and/or open environments*[48] and (b) their ability to engage in *autonomous (lethal) decision-making,*[49] *autonomous (lethal) targeting*[50] and *autonomous (lethal) force*;[51] (2) they could be used as *offensive autonomous weapons*;[52] (3) these are all part of advances in *Artificial Intelligence* (AI) that will distinguish fully autonomous weapon systems from the existing weapon technologies[53] because true AWS will be able to operate without human oversight, instead – on the basis of "artificial intelligence algorithms," also potentially permitting them to engage in "machine learning."[54] The detailed discussion of related artificial intelligence or machine learning as distinct phenomena or processes is beyond the scope of this article, which aims to refer to them in the context of the dilemma analysis.

Dilemma analysis

The debate on and around the just-delimited robot weapon category of AWS, as already specified above, concerns primarily the possibility of their autonomous lethal force. Those opposing AWS deployment and calling for an international pre-emptive ban on AWS build their argument on the basis of the deep convergence of a deontological view-

point and a consequentialist standpoint.[55] With regard to the former, by "taking ethical conduct by humans for granted ('humans are ethical, and robots are not')," [56] they believe, as accurately summarised by Peter Asaro, that it is immoral by itself to kill without the involvement of human reason, judgement and compassion and outsourcing lethal decisions to machines may automatically mean the regress in ethics and morality, thus, it should be illegal.[57] With regard to the latter, their deontological position is supported by the multi-dimensional consequentialist analysis of "expected benefits and costs flowing from AWS deployment," which they use to substantially prove costs are likely to "outweigh the sum of the expected benefits."[58] This is where their argument meets the multi-dimensional counter-argument by those "giving up on human morality altogether ('humans fail to act ethically, so we need ethical robots')"[59] and insisting that robots's potential ability to perform better than humans in the battlefield means the achievability of ethical robot autonomy[60] and establishes a "moral imperative" to make use of AWS in combat.[61] The following paragraphs serve to provide a deep insight into these multiple argument – counter-argument chains in relation to AWS deployment. Although the primary issue of the dispute largely drives their content, the argumentation goes deeper in many regards touching on all potential implications of AWS deployment, as either directly or indirectly related to the central concern.

Dilemma no. 1: (Un)predictability of AWS performance

The prospects of exercising human control over as well as being able to understand and predict the patterns of AWS performance is a core concern underlying multiple issues related to other dilemmas and their (sub)sections.

The argument against AWS deployment warns against the difficulty of reliably predicting the behaviour of complex autonomous systems.[62] Humanity risks not only having "little knowledge of — or control over — what is being done in its name,"[63] but also potentially facing multiple challenges and dangers, including the one of robots "running amok."[64]

Double-edge sword of pre-programming

A mathematical formula or a magic algorithm for intelligence does not exist[65] meaning a machine will never be fully identical to a human. While the importance of human judgement and reason by themselves

versus computational formulas will be highlighted within the next dilemma, here will be presented the analysis of the technical aspect of the inability to fully replicate human intelligence through algorithms. With regard to this, there is the choice between the two contradictory pre-programming options for AWS, neither of which seems attractive.

a. *Software rigidity.* Robots lack "situational awareness,"[66] "contextual intelligence or common sense, on par with humans."[67] That means decisions implemented via an autonomous system cannot be based on observations of the situation to which the decision relates but are "based on whatever information is available through experience and foresight at the time the machine is programmed."[68] Robots "cannot be programmed for all eventualities" though, especially in military scenarios,[69] and even "sophisticated algorithms are subject to failure if they face situations outside their intended design parameters."[70] This may result in contextual misperformance caused mainly by the two major challenges in this regard: the "problem of relevance" of information, and the "problem of representation" of subjects and objects in combat situations.[71]

   Firstly, one only relies on relevant information in a given context that is "relatively easy for humans to do, but very difficult for computers."[72] The latter may potentially face the challenge of information or data limitation combined with their limited ability to capture subjective human meanings.[73] Secondly, testing environments "may be substantially different than more complex, unstructured, and dynamic battlefield conditions"[74] and there is a risk that a robot's world model "may not correspond exactly to reality" due to the limitations of its sensors and processing algorithms, harsh conditions such as dust, noisy and low-light conditions, dynamic environments or explosions, which may drastically change the environment.[75] Looking ahead, "rigidity can easily lead to bad consequences when events and situations unforeseen or insufficiently imagined by the programmers occur, causing the robot to perform badly or simply do horrible things."[76]

b. *Software flexibility.* The agenda aimed at the creation of (artificial) super-intelligence,[77] alternatively – strong artificial intelligence[78] or superhuman intelligence,[79] is also a concern. The idea behind is to develop strong machine intelligence capable of reaching

and potentially surpassing, outstripping or exceeding human intelligence.[80] AWS super-intelligence promises to be "capable of independently interpreting and even setting its goals and acting to attain them," acquiring "great volumes of data themselves and categorise it in new, sometimes unexpected, ways," acting "on that information with speed and precision unobtainable by human controlled systems" and "capable of learning from experience and improving performance."[81] The latter implies the process of machine learning, which means a robot will act by rules that are "not fixed during the production process, but can be changed during the operation of the machine, by the machine itself."[82] The potential capabilities for "shape-shifting" in reconfigurable systems[83] and creating other robots through "self-replicating"[84] have also been mentioned as potential components of the overall picture. In the era of super-intelligence it may be too hard to foresee the behaviour of a robot "introduced to novel situations"[85] or to "predict with reasonable certainty what the robot will learn."[86] AWS warfare risks evolving "not only beyond human control," but even possibly "beyond human understanding."[87] As a potential outcome, super-intelligent machines "may pose a threat to humans, either deliberately in pursuit of its own goal or inadvertently in optimising some pre-set goals."[88] At best it may lead them to overwrite their own programming, especially with regard to the most fundamental aspects of the Laws of War (LOW) and Rules of Engagement (ROE),[89] and at worst humanity may face a "robot revolution."[90]

Inherent software unpredictabilities

The "inherent weaknesses in AWS" encompass those mechanical issues that "will always be the Achilles' heel of any tasking and deployment of any weapon system."[91]

a. *Software imperfections.* All programs have "bugs" implying "errors in the logic of the program itself" that are typically undetectable and may either manifest themselves in specific circumstances, usually only during the execution of the program, or even never manifest themselves.[92] Additionally, any system is subject to breakdowns, malfunctions, glitches.[93] The computer program used in the robot's on-board computer may consist of "millions

of lines of code" written by teams of programmers, none of whom knows the entire program, that results in the impossibility to "predict the effect of a given command with absolute certainty, since portions of large programs may interact in unexpected, untested ways."[94] To highlight, "as complexity of any system increases, the more opportunities exist for errors to be introduced."[95] Programming bugs and system malfunctions lead to accidents[96] and "mistakes by military robots may be fatal."[97]

b. *Cybervulnerability.* Any system is subject to interferences.[98] Bugs are typically considered to be software vulnerabilities that can be exploited by hackers to cause the system to do something other than what it is designed to do on a regular basis, or even can lead to AWS being hijacked.[99] Unsurprisingly, as computer programs become more sophisticated, they simultaneously become more vulnerable to cyberattacks.[100] There is a risk that using this channel, the enemy – be it a state or a non-state actor – "might be able to use cyber means to take control of an autonomous weapon system and direct it against friendly forces or a civilian population."[101]

Superhuman pace of battle

AWS will be "able to process information and reach decisions sequentially and via parallel processing at speeds that are orders of magnitude faster than humans."[102] They will be able to make decisions in nanoseconds, while humans may need a minimum of hundreds of milliseconds for the same.[103] Such a pace of the battle, where decisions are taken with "superhuman speed"[104] meaning "the speed of action on the battlefield would eclipse the speed of human decision-making,"[105] may be "way beyond the speed of human intervention" leaving humans with "little control over the battlespace."[106]

Unpredictability of device-device interactions

a. *Coordinated attacks.* Increasingly, "it will become necessary to deploy multiple robots to accomplish dangerous and complex tasks" in a form of multiple robot system architectures executing coordinated attacks.[107] One potential drawback of a network architecture involving autonomous devices may be that, because of the complexity of such a system, the interaction can be unpredictable.[108]

b. *Friendly-hostile interaction.* As more and more countries (attempt to) develop AWS and autonomous counter defences, "these weapons as well as command and control systems will inevitably interact" and when "any mobile device controlled by software programs interacts with a competing hostile device controlled by unknown software, the result of the interaction is scientifically impossible to predict."[109] The "speed of their unpredictable interaction" may further exacerbate this concern.[110]

The counter-argument rests on two pillars with the first one critisizing biased framing of AWS (un)predictability, and the second one clarifying that "there is no such thing as 'complete' autonomy in the sense of a machine operating entirely independently of any human."[111]

Biased framing of AWS (un)predictability versus conventional warfare

a. *Unpredictability of existing systems.* The problem of malfunction is not unique to AWS but is the case with different weapon systems ranging from catapults to more complex computer attack systems.[112] Cyberattacks are also not new, as has been repeatedly demonstrated in cases with predator drones that "have been regularly hacked by militants."[113]
b. *Unpredictability of human conduct.* Human soldiers, in turn, are subject to a number of psychological factors, which are to be discussed in detail within the context of the next dilemma, that "render their behavior unpredictable."[114] In turn, software-based AWS may "potentially remove much of the unpredictability of human behavior in the battlespace."[115] By taking the human being out of the decision-making chain, AWS may at least "partially decouple the limits of the system from the limits of its operators."[116]

Illusion of unrestrained robotic autonomy

It is fundamentally incorrect to describe autonomous systems "as being 'independent' machines that operate 'without human control' " or as " 'intelligent" machines having the capacity for 'choice' or 'truly autonomous' operation."[117]

a. *AWS is never human-free.* A "fully autonomous system is never completely human-free."[118] Fundamentally, the "development of an artificially intelligent system is in fact just an exercise in software development," where subsequently the only function of a computer is to run the installed software, and although it may seem that "the system itself is 'choosing' between two alternative courses of action," in fact, the choice is "made in advance" by the person writing the program.[119] Whether concrete actions are "explicitly programmed into a machine," or whether "technologies of artificial intelligence are employed to allow the machine to adapt its behaviour dynamically," in either case AWS "behaviour originates not in the machines themselves, but in the minds of their developers." [120] The only difference between these actions and more familiar actions in the battlefield is that there will be a apronounced "lag-time between the latent human decisions built into the causal architecture of the weapons system itself and the anticipated combat effect of that weapon system that later eventuates."[121]

b. *AWS are never order-free.* Additionally, AWS "autonomy should be considered in light of the existing command and control structure" that does not presuppose operation without orders.[122] Human combatants are in fact expected to act in accordance with "a regulatory and governance framework ranging over a set of considerations, from the international law of the sea, to humanitarian law and a range of treaty obligations, all the way to specific rules of engagement," and full autonomy to act without external restraint has seldom been granted even to the human commanders.[123]

Dilemma no. 2: Dehumanization of Lethal Decision-Making

The removal of a soldier from the battlefield leads to "dehumanization of killing" that may already, to some extent, manifest itself in the use of remotely operated drones.[124] AWS may simply mean a step further in dehumanization of warfare.[125]

The argument against AWS deployment insists that further robotisation may transform warfare into "unempathic automated industrial process."[126]

Combat deprived of healthy human emotion

Humans tend to engage in emotional reasoning[127] and human emotions may play a positive role in combat.[128] "Healthy" emotions,[129] which may include an innate reluctance or inhibition to killing, guilt, concern, mercy, the ability to empathize and the capacity for compassion,[130] serve as "drivers of prosocial behaviour and moral sensitivity" producing "a major obstacle to killing in war"[131] and "an important check on the killing of civilians."[132]

AWS run by a program has no human emotions, thus, will be unable to employ them.[133] Not only will the deployment of these so-called soulless killers[134] in combat "make killing easier,"[135] but it will also result in "the deprivation of hope" for some kind of empathy, mercy, and reprieve.[136]

Combat deprived of human judgement and reason

As "the context gets more complex, it becomes impossible to anticipate all the situations that soldiers will encounter, thus leaving the choice of behavior in many situations up to the best judgment of the soldier."[137] Human decisions in combat are guided by human judgement and human reason.[138] The significance of those in the military context cannot be denied. Let alone that the boundaries between groups such as "friend" or "foe" are "often poorly defined and heavily value-laden,"[139] recognising a civilian and a combatant is of central concern. "This distinction makes it legally permissible, at least sometimes, for combatants to kill enemy combatants" and makes it "almost never legally justified for combatants to kill innocent civilians," however, at the same time "combatants retain certain rights, like the right to surrender, and not to be killed unnecessarily" and there are "cases in which it is legally permissible to kill civilians."[140] The distinction is blurry. In addition to "the lack of a clear definition of civilian,"[141] with regard to combatants it is also "not just a matter of uniform; soldiers who are wounded, have surrendered or are mentally ill are also immune."[142] This distinction may also be highly problematic "in guerrilla and insurgent warfare, in which combatants pose as civilians."[143] It may be challenging for robot sensors not only "to distinguish between a man carrying an AK-47 and a man carrying a walking stick,"[144] but especially to distinguish "between a civilian carrying a weapon and a combatant."[145]

To specify on their role, in the military context, human judgement and human reason are, firstly, "necessary to comply with the *law*."[146]

Law is by its essential nature imperfect, incomplete, and subject to interpretation.[147] Human situational understanding and judgement, which enable considering and drawing insights from different, potentially incompatible or contradictory, perspectives thereby keeping the legal system on track, "exceed any conceivable system of fixed rules or any computational system."[148] Secondly, there may be the "distinction between fundamental *morality* and practical law."[149] The "ability to think morally based on one's values, and to give oneself the moral commands" is also a "distinctive human characteristic."[150] To put it bluntly, moral reasoning also cannot be codified or programmed.[151] "None of these are fixed values,"[152] and human reasoning often involves "qualitative rather than quantitative judgements."[153] Depriving the combat of human judgement and human reason will eliminate the human determination of morality and legitimacy of lethal force,[154] including the right to surrender.[155] Importantly, the human ability to disobey illegal and immoral orders will also be eliminated.[156]

Combat deprived of military honour

One interesting argument against AWS warfare is that humans are capable of "morally praiseworthy and supererogatory behaviour," as exemplified by heroism or "going beyond the call of duty," something that machines will not be able to replicate.[157] This is linked to the concept of military honor, which is a human value because robots will follow orders without being aware of making sacrifices.[158] Using AWS in combat will be a violation of this principle.[159]

The counter-argument denies "the mere potentiality" of a human combatant's mercy, compassion or honor "should make a difference if, in fact, this potentiality does not materialize."[160]

Biased framing of AWS warfare versus human warfare

It is necessary to "beware of idealizations of human warfare."[161] To begin with, as Ronald Arkin quoting Immanuel Kant, Albert Einstein and Sigmund Freud summarised, war and aggressive tendencies seem to be ingrained in human nature because humanity's "propensity to wage war has gone unabated for as long as history has been recorded."[162] On top of that, humanity also "has a rather dismal record in ethical behavior in the battlefield,"[163] which may, to some extents, result from certain "performance-hindering conditions."[164]

a. *Human biological factors.* Biological limitations to human effective and ethical performance in the battlefield may include the requirement for breathable air, rest and sleep, drinkable water and food, as well as the physical extremes of acceleration and cognitive load, and also the vulnerability to temperatures, radiation, biological and chemical weapons.[165]
b. *Human emotional-psychological factors.* "While it is certainly correct that emotions can restrain humans, it is equally true that emotions can unleash the basest of instincts."[166] Negative emotions or psychological factors can enter the scene.[167] Emotional distortions can occur.[168] Frustration, fear, stress, hysteria, panic, spite, hatred, anger, hate, prejudice, revenge, vengefulness, resentment, mental disturbance or trauma, as well as self-preservation,[169] and importantly, human lack of an "offensive spirit" in certain circumstances[170] – are all part of the list of factors that may potentially "cloud" human judgment.[171]
c. *The fog of war.* The "fog of war"[172] or the "turmoil of war"[173] may additionally hinder effective human performance because, in the military context, interactions often have to be carried out in noisy, stressful, and confusing conditions and are additionally challenged by the pressures of time, environmental hazards, degradation of communications, multiple control problems and perceptual challenges, as well as decisions sometimes have to be made with unclear orders or contradictory information in stressful situations.[174]

Less inhuman AWS warfare

The factors mentioned above serve as potential explanations for "human error"[175] in the forms of human underreaction and overreaction.[176] While the former may result in prolonging the war,[177] the latter may drive excessive and indiscriminate uses of force, contribute to war crimes, friendly fire incidents and/or unjustified collateral damage, including noncombatant casualties and damage to civilian property.[178]

As AWS will be "devoid of negative human emotions,"[179] "resilient to adverse psychological effects that underlie the perpetration of some unlawful acts by human actors,"[180] "immune" to other human "performance-hindering conditions,"[181] and will be able to reduce the negative impact of the "fog of war,"[182] partially through removing the need for vulnerable control and communication links,[183] they have "the poten-

tial to ultimately save human lives (both civilian and military) in armed conflicts."[184] If "programmed to never break the laws of war," AWS would be "incapable of doing so."[185] "A notion proposed by the proponents of lethal autonomous robots" is that AWS strict reliance on preset technological "fixes"[186] and "data-driven, bias-free analysis"[187] will allow AWS both to eliminate moments of hesitation or mercy when killings are objectively necessary for ending the war sooner, in turn saving many lives *overall*,[188] and to put an end on deliberate violations of the laws of armed conflict,[189] in turn promising "fewer war crimes, fewer civilian casualties."[190] That means AWS may make war "less inhumane through lessening the human element from warfare."[191]

In addition, as human soldiers and autonomous weapon systems may be deployed in integrated architectures in the future warfare, the potential capability of AWS to independently and objectively monitor and report (un) ethical behavior in the battlefield by all parties may lead "to a reduction in human ethical infractions."[192]

A note on military honor

Although, as indicated above, robots may be blamed to be unaware of making sacrifices and unable to replicate human heroism,[193] the counter-argument is that, in real combat, only a few combatants may seek combat glory, while roughly ninety-nine percent of them simply want to complete the mission efficiently and with the least possible amount of casualties.[194] Importantly still, many medals for heroism are awarded for defensive actions, and AWS may actually be ideally suited for the overall defensive posture thereby compensating for the potential lack of human military honour in AWS warfare.[195]

Dilemma no. 3: Depersonalisation of enemy (non-)combatant

The dehumanizing effect of AWS may be susceptible to other problems. That is "depersonalization of war"[196] made possible through the combination of "depersonalized forms of responsibility,"[197] which are to be discussed as a part of the legal challenges associates with AWS warfare, and "*depersonalisation of the enemy*."[198]

The argument against AWS deployment regarding the latter component is built on the assumption that by following analogies of the dronification of military interventions, the use of lethal robots will further depersonalise war and methods of killing by removing all

human attributes from the representation of the enemy and turning enemy (non-) combatants into objects deprived of moral value.[199] This practice of "objectivisation" may turn warfare into "a factory of death."[200]

Absense of inter-personal relationship[201]

Lethal force "has always been an intensely personal affair" with a human being physically present at the moment of the release of force and taking this decision.[202] The practice of "killing at a distance"[203] brought about by remote-controlled systems will be "taken to a next level through the introduction of the autonomous release of force."[204] AWS threaten to increase both distancing and detachment.[205] While the physical distance from the act of killing may not be greater, the psychological distance will no longer play a significant role in AWS warfare.[206] This not only will render enemy (non-) combatants "less visible"[207] through reducing them to "targets" in a "dislocated reality,"[208] but also will exacerbate "moral disengagement" of humans from lethal decisions in combat.[209] By making it "significantly easier for them to make the decision to kill,"[210] the deployment of AWS by humans may lead to more killing,[211] and even encourage more unethical choices.[212]

Automated death

Automating death by "algorithm"[213] means treating enemy (non-) combatants simply as "things thrown out of the realm of good and evil"[214] or objects "eligible for mechanized targeting."[215] In combat, where situational decisions made by individual human combatants will be replaced with general choices made by people defining the behaviour of AWS in advance,[216] "the generality of the decisions" will dominate decision-making dynamics.[217] However, in sample distinctions between combatants and civilians, there are "shades of grey" as combatants retain certain rights, including the right to surrender and not to be killed unnecessarily, and as it is legally permissible to kill civilians in certain cases.[218] Machines "missing battlefield awareness or common sense reasoning to assist in discrimination decisions"[219] may potentially leave "behind them a hecatomb of innocent victims."[220] This represents a threat to the fundamental values of human dignity and human life.[221]

The counter-argument is based on the assumptions that the "introduction of AWS does not mean the introduction of an altogether new

quality of warfare"[222] and that the deeper unbiased analysis may actually reveal that abandoning AWS may "deny protections to civilians and soldiers."[223]

Biased framing of AWS warfare versus conventional warfare

Anzhelika Solovyeva
Nik Hynek

a. *Conventional warfare.* Largely, the features that make AWS problematic with regard to the values of human life and dignity may, to some extents, be present in conventional acts of war.[224] Firstly, "much of war is mechanical slaughter" and much of "modern warfare is impersonal killing at a distance,"[225] as has been experienced through the use of, for example, over horizon weapons, indirect fire, or buried improvised explosive devices in combat.[226] Secondly, historical experience may provide a plenty of examples of war practices that have been clear cases of war crimes and violations of the human dignity, meaning that not weapons themselves but, rather, uses to which they are put may potentially be contrary to the values of human dignity and human life.[227]
b. *Manner of death.* First and foremost, "seeing the man's eyes as he stabs you doesn't make your death any more palatable."[228] For victims whose life and dignity are at stake, "it is a matter of indifference whether the threat they are exposed to comes from manned or unmanned weapons, provided all other parameters of the situation are equal."[229] The manner of death will basically be no different in the age of AWS warfare because there is nothing more dignified in, for example, being mowed down by a machine gun or blasted to bits by a bomb, burning alive in an explosion or slowly suffocating from a sucking chest wound.[230]

Reduction of total war casualties in AWS warfare

At this point, it is worth recalling an argument that "removing the human element from the equation could be potentially beneficial."[231] Additionally, "increased accuracy saves lives" and, as emphasized by Ronald Arkin, AWS will be the next-generation, precision-guided munitions.[232] Due to AWS being "more accurate in their targeting and more considerate in their fighting habits than manned systems,"[233] "the increased depersonalization in the deployment of force brought about by AWS may thus lead to *greater personalization* in targeting outcomes and saving lives or preventing unwarranted injuries."[234]

Dilemma no. 4: Human-machine nexus in coordinated operation

Even if deployed, AWS "will not, at least initially, entirely replace human soldiers,"[235] but will rather be "integrated into human warfare."[236] That means while their numbers are expected to decrease on the AWS deploying side, a portion of human soldiers will "fight alongside AWS" because an army of robots fulfilling all or a large majority of functions in an armed conflict is not likely in the near future.[237]

The argument against AWS deployment asserts that in such a configuration "human beings will start to be placed in harm's way as a result of the operations of robots."[238]

Illusion of «push-button» war

By falling into an "illusion that war can be fought without casualties"[239] in a form of a "push button" war implying "the enemy is killed at a distance, without any immediate risk to oneself,"[240] humans "can fall victim to automation bias, trusting too much in the machine."[241]

a. *War initiation.* The availability of AWS "may mean that military conflicts are initiated with the intention that they can be completed without placing warfighters in harm's way" but, in reality, there is a high chance that human warfighters "may find themselves involved in conflicts" either because a weapon system may fail due to its software unreliability or because winning a victory may turn out to be beyond the capabilities of AWS due to changed circumstances.[242] The letter may involve an enemy action, or the operation being ill-conceived in the first place.[243]
b. *Military tasks and operations.* Similarly, as "robots can play a useful role in military operations, warfighters will rely on them to complete the tasks to which they have been assigned," but the possibility of robots' failure may not be excluded and, in case it happens, "human lives may be placed at risk."[244]

Challenges of co-existence in coordinated operation

a. *Priorisation of AWS value over human life value.* A combination of one robot's price that "may range from $100,000 to millions of dollars in cost,"[245] AWS significant "military utility,"[246] which is to be discussed below in this article, as well as "valuable intelligence" it carries implicitly means fellow soldiers' lives may be placed "at

risk in order to defend, service, or recover" it upon necessity.[247] The other side of the coin may be the anthropomorphised image of AWS leading soldiers to "often treat them as fellow warriors" and being sometimes "prepared to risk their own lives to save them."[248]

b. *Multiple control challenges.* Even when the abundant causes of the unpredictability of AWS performance, which potentially stand also behind multiple control problems, are left on the side, there still remains a related challenge, but of a different nature. While both manned and unmanned components may be expected to operate in conjunction, including with battlefield surveillance devices,[249] and while increasingly multiple robots may be deployed in complex tasks and missions it is still unclear how the "proper architecture for control" should look like.[250]
c. *Friendly fire.* The boundaries between "friend" and "foe" groups are "poorly defined" and they are "heavily value-laden."[251] Thus, reliably pre-programming these identification parameters is challenging in the first place. A robot "that cannot distinguish between targets may be highly prone to friendly fire incidents."[252] Even if turning these values into an algorithm is possible, AWS may still suffer from inherent software weaknesses and unpredictabilities potentially leading fellow soldiers to be "accidentally killed by machines."[253]
d. *Order refusal.* Human military conduct "entails making judgments with imperfect knowledge in complex, ambiguous and dynamic situations," but AWS will be "ill-equipped" to comprehend phychological and subjective phenomena as well as dynamic goals.[254] That implicitly means a "conflict may arise" between some pre-programmed instruction and real combat demands leading to the refusal of an otherwise-legitimate order by the machine.[255]
e. *Negative impact on squad cohesion.* Robots equipped with video cameras and sensors to record and report soldiers' actions in the battlefield may "negatively impact the cohesion among team or squad members by eroding trust with the robot as well as among fellow soldiers who then may or may not support each other as much anymore, knowing that they are being watched."[256]

The counter-argument rests on the assumption that the "primary rationale for the development of unmanned systems in general (re-

mote-controlled weapons and AWS) is their ability to protect personnel who are kept out of harm's way."[257] By "removing soldiers from the most dangerous and life-threatening missions,"[258] robots will facilitate "a reduction in friendly casualties."[259]

Removal of human soldiers and operators from risks of war

Using either unmanned or autonomous technology means that an army has "no skin in the game."[260] However, as mentioned above in the context of discussing the dimensions of distance, by increasingly removing fellow soldiers from the risks of war not only *physically*, but also *psychologically*, AWS will significantly "reduce the potential cognitive overload of operators and supervisors."[261] Overall, beyond increasingly ensuring physical safety, AWS may increasingly contribute to the reduction of psychiatric damage or trauma, and even psychiatric casualties mainly linked to the suicide practice, among active duty friendly forces.[262]

Minimisation of harm to fellow soldiers in coordinated operation

a. *Outsourcing tasks.* Although AWS will not necessarily replace humans in combat, they may at least "reduce their exposure to life threatening tasks"[263] as machines can perform dull, dirty, and dangerous tasks and missions that human combatants may prefer to avoid.[264]
b. *Friendly fire bias.* As already pointed out previously, human soldiers "can become emotionally disturbed, suffer from battle fatigue, or simply decide to act outside of the chain of command," which can lead, among others, to "friendly fire incidents."[265] AWS designed to remove human cognitive shortcomings[266] and psychological shortcomings[267] in decision-making promise "fewer friendly fire incidents."[268]

Dilemma no. 5: Strategic considerations

The argument against AWS deployment in this regard is two-fold and rests on the national and the systemic levels of analysis. It is believed that artificial intelligence in the military "will, in the very near future, have a profound impact on the conduct of strategy and will be disruptive of existing power balances."[269]

National strategic risks

As super-intelligence may "evolve beyond human understanding and control,"[270] there is a danger of the loss of human control "over war" initiating, escalation and termination.[271]

a. *Loss of control over one's own national strategy.* There may arise the so-called "Strategic Robot Problem,"[272] which Srđan T. Korać summarised as "the possibility of loss of human control over the conduct of military operations, even the entire war, should we equip robots with artificial intelligence to decide independently on strategic, operational and tactical levels."[273] Let alone the risks related to software coding errors or malfunctions, and especially cyberattacks,[274] when "decisions are made with inhuman speed, the potential for events to spiral out of control is obvious."[275] That makes strategy in a world with autonomous weapons "impossible to predict."[276]
b. *Loss of understanding of one's own national strategy.* Human strategy entails the instrumental use of violence in the pursuit of goals, usually social goals, has psychological attributes and a cultural dimension meaning human strategic goals may be hard to measure, and is essentially dynamic meaning human strategic goals may change in response to emerging situations and opportunities.[277] In turn, AWS will be "ill-equipped to gauge these subjectively experienced and dynamic goals compared to more readily quantifiable goals," implying their limited ability to capture and reproduce subjective meanings inherent in human strategy. [278]

Systemic strategic risks

By falling into an illusion of "a risk-free war," humanity may underestimate the potential structural risks.[279] The combination of factors is listed below to support the assumption that the militarization of artificial intelligence may not only "create significant problems for the stability of the international system,"[280] but may also pose a serious threat to the "ability of international bodies to manage conflicts."[281] An exacerbating factor in this regard, whose degree of manifestation is positively linked to the systemic strategic risks discussed below, is the increasingly direct incorporation of "cyberwarfare (along with its lower-threshold counterparts of cybercrime and cyberterrorism) into

armed conflict in the physical world" as an outcome of AWS development and deployment.[282]

a. *Proliferation and strategic competition.* As soon as "one nation is capable of deploying AWS that can operate without human oversight then all nations will have a powerful incentive to do so."[283] This may provoke a new strategic competition between major powers and lesser powers, and considering the nature of autonomous weapons technology this arms race will most likely be global in scope.[284] Driven by "dual-use" technologies of artificial intelligence and robotics[285] and additionally, unlike nuclear weapons, requiring no costly or hard-to-obtain raw materials, autonomous weapons may become "the Kalashnikovs of tomorrow."[286] Proliferation of AWS may "occur via exports, including to the grey and black markets"[287] or as a result of some states developing their own AWS technology.[288]

   The potential acquisition of AWS by non-state actors is also a concern[289] where beyond "further privatisation of violence on the global level by increasing the capacity of private military companies,"[290] it will only be a matter of time until this technology falls into the hands of terrorists, criminal cartels, and extremist groups. [291]

   The concomitant danger is that proliferation may proceed without the expected level of safeguards.[292]

b. *Lowered threshold and normalization of armed conflict.* Political costs of war "come with wartime casualties" and casualties are a significant reason of armed conflicts not being more common.[293] In AWS warfare, the "political calculus would not have to take into account the number of fallen soldiers,"[294] while "expendable" robots may be "risked in provocative adventures."[295] As this potentially promises "easier internal legitimisation and execution of military interventions,"[296] it may result in lowering of the threshold for armed conflict,[297] and "armed conflict no longer being a measure of last resort."[298]

   At the same time, as sending machines to war does not exact "physical and emotional toll on a population," the national public of AWS-equipped states "may over time become increasingly disengaged and leave the decision to use force as a largely financial or diplomatic question."[299] This will produce the "normalization" of

armed conflict.[300] Merging lethal robot technology and private entrepreneurship in meeting military demands may further contribute to removing "low intensity wars outside of the public eye."[301]

c. *Accidental and non-attributable war.* There will significantly increase the risk of "an accidental war being triggered by the decisions of *one or more* autonomous weapon systems."[302] In AWS warfare, supersonic or hypersonic (defence) systems of one state will interact with equally fast systems belonging to another state and the "speed of their unpredictable interaction" may potentially "trigger unintended armed conflicts before humans had the opportunity to react."[303] In an asymmetric war, it may mean the significantly decreased amount of time available for the other side to determine whether an attack is imminent or under way, and how to respond.[304] Misinterpretation may invite pre-emption and undesired escalation.[305] This places states under the pressure to mobilize their forces that further increases the chance of a war occuring in error.[306]

   To dig deeper, the probability of "unintended initiation or escalation of conflicts outside of direct human control" in AWS warfare,[307] exacerbated by the furtile environment of "the anonymity of cyberspace,"[308] may make possible "shielding" human perpetrators from the responsibility for "what might have otherwise been considered a war crime."[309]

d. *Facilitation of asymmetric warfare.*[310] AWS army is "the product of a rich and elaborate economy."[311] The "imbalanced system of haves and have-nots" in relation to autonomous weapons[312] will mean "the completely asymmetric 'push-button' war,"[313] in which "deadly robots may in some cases be pitted against people on foot."[314] This raises a question whether it still makes sense to talk about "war," "as opposed to one-sided killing."[315]

   The use of AWS may encourage retaliation and reprisals by the other side.[316] The counter-actions in this asymmetric war may include terrorism at home and abroad,[317] potentially with civilians of AWS-deploying states being "the next-best legitimate targets,"[318] as well as intensifying efforts to acquire nuclear or biochemical weapons.[319] The potential use of cyber means by the enemy or non-state actors "to take control of an autonomous weapon system and direct it against friendly forces or a civilian population"[320] may also be a part of this equation.

e. *Challenges of post-conflict reconciliation.* The possibility of a lasting peace after an armed conflict requires "diplomacy and human relationships that machines would not be capable of delivering" meaning AWS warfare may make "peaceful reconciliation most difficult to achieve."[321]

The counter-argument insists that autonomous weapon systems may be strategically, tactically and operationally beneficial to the conduct of strategy and may potentially make war less brutal at least or render inter-state war obsolete at best.

Qualitative improvement of national strategy

With regard to the issue of human control and understanding of one's own national strategy, in the first place, assuming that AWS will "supplement, not replace, human combat forces,"[322] strategy that involves humans, no matter that they are assisted by AWS in the battlefield, "will retain its inevitable human flavour."[323] Potentially still, as there is the probability that people "can deviate from orders" while autonomous systems "will do precisely what they are programmed to do," the deployment of the latter may even potentially result in the "increased leaders' control over how their forces behave in crises."[324] Going beyond, due to the potential performance superiority of autonomous weapon systems in comparison to human combatants and remotely operated systems in combat,[325] AWS will even be "able to improve the quality of human decision-making at strategic levels" as well as will bring "tactical and operational advances."[326]

a. *Force multiplication.* Robots may bring the potential for force multiplication in military deployments.[327] This may materialise both through each robot "effectively doing the work of many human soldiers"[328] and at the same time through "allowing fewer personnel to do more."[329] The latter implies that "one soldier on the battlefield can be a nexus for initiating a large-scale robot attack from the ground and the air."[330]
b. *Expanding physical limitations.* Firstly, AWS better-informed and faster reaction needs to be highlighted[331] because AWS abilities are greater than those of humans with regard to data absorption and data analysis.[332] Not only will AWS be able to observe a large number of relevant aspects due to their superior sensor

abilities[333] and to master huge amounts of data, potentially recognising patterns that may otherwise be missed,[334] but they will also be "much faster at processing enormous amounts of data"[335] and will have quicker potential reaction or response times than the best human could have.[336] Additionally, if programmed to do so, AWS may more effectively "learn from its mistakes, and improve its algorithms as the conflict goes on, while humans remain rooted in their entrenched cognitive heuristics and groupthink."[337] Secondly, AWS may extend a warfighter's reach by enabling military forces "to reach deeper into the battle space by, for example, seeing or striking farther;" and thirdly, AWS may expand the battlespace by allowing combat "to be conducted over larger areas than was previously possible."[338]

c. *Control and communication links rendered obsolete.* The speed of human decision-making in combat may be "further slowed down through the inevitable time-lag of global communications" between human operators and remote-controlled systems.[339] These links may also be threatened by electronic counter-measures by hostile forces, different environmental factors and other exigencies of the "fog of war".[340] AWS will "render constant control and communication links obsolete"[341] and their ability to operate in the absence of these links is an obvious military advantage.[342]
d. *Reduced political costs and democratic resistance.* Machines are "expendable" because "their loss does not cause emotional pain or political backlash."[343] The possibility of sending "an army of machines to war — rather than friends and relatives,"[344] may "remove the democratic resistance to military deployment" in case of its necessity under a variety of circumstances.[345]

Positive systemic strategic impact

a. *Less brutal war at least.* Referring to the work of Ronald Arkin, George R. Lucas summarised that the development and use of autonomous robotic technology, may at least "render war itself, and the conduct of armed hostilities, less destructive, risky, and indiscriminate."[346] In fact, as already pointed out previously, having "robots fight for us promises to dramatically reduce casualties on our side"[347] that may also be true with regard to the

total number of human casualties in war due to AWS "accurate determinations."[348] The general destruction associated with armed conflict may also be expected to decrease with potentially more precise machines.[349]

b. *Potential absense of war at best.* More ambitiously, the use of AWS, or potentially superintelligence, "would represent a chance for a world without armed conflict."[350] Firstly, as AWS will be "ill-equipped" to comprehend psychological and subjectively experienced phenomena underlying human strategy,[351] the wider deployment of this warfare technology may "assist humanity in transcending some of the causes of armed conflict – be they cultural or material."[352] Secondly, in ideal case, although it has been highly debated, AWS may make possible a war in which the sides send only robots to do the fighting, and each party to a conflict can "only inflict economic damage" on the enemy[353] and similarly "carries no existential risk, and bears no cost beyond the economic."[354] If this is to become a reality, "war will cease to be a desirable option by nation-states as a means of resolving their differences."[355]

Dilemma no. 6: AWS operation in law(less) zone

The argument against AWS deployment insists that they "will operate in a lawless zone."[356] The use of military robotics has been objected on the grounds that it may make easier the decision to initiate a war, in an apparent violation of jus ad bellum.[357] What is more, "the technical ability to properly discriminate against targets, as required by jus in bello," has also been a notable concern.[358] Both points are amplified by the problem of attribution of criminal responsibility.[359]

Threat to (non-)use of force norm

"During the larger part of the last two centuries," international law has been developing "to constrain armed conflict and the use of force" and to make them the options of the last resort.[360] When they ratified the UN Charter, states have agreed not to use the force without the permission given by the United Nations, except for defensive purposes.[361] AWS, especially considering their potentially unpredictable nature and the subsequent responsibility gap – the latter discussed below, may make this norm against the use of force, which has been paramount in ensuring global security, breakable.[362]

(Non-)compliance with international humanitarian and human rights law

AWS may bring "significant obstacles to complying with international humanitarian and human rights law."[363] The concept of human dignity and the right to life lie at the heart of international human rights law.[364] The right to *dignity* and the right to *life* form the "protect life" principle that is "the guiding star whenever lethal force is used."[365] However, in the exceptional circumstances that prevail during armed conflicts, "human rights law remains valid, but it is interpreted with reference to the rules of international humanitarian law" (IHL).[366] Concerns with regard to the latter "have so far most often related to the legal principles of distinction and proportionality,"[367] to specify, how to program AWS to act in such a way that the principles of discrimination and proportionality, as demanded by the law of international armed conflict, are applicable in the battlefield in the age of AWS warfare.[368] These principles "reflect the tension between these opposite goals" in combat, where the former "embodies the necessity of differentiating military personnel and militarily significant targets from civilians and civilian object" and the latter "embodies the requirement that any attack which could have adverse consequences for civilians must have a military objective which is not excessive with regard to the potential civilian harm."[369] As IHL suffers from terminological hurdles and obscurities, most importantly "the lack of a clear definition of civilian,"[370] as well as the problem of "contradictory or vague imperatives,"[371] its interpretation is understood to involve "subjective estimates of value and context-specificity" and human judgement.[372] AWS "restricted abilities to interpret context and to make value-based calculations"[373] may imply its inability to comply with international humanitarian law.[374]

Legal review challenge

"The obligation to carry out legal reviews of new weapons under article 36 of Additional Protocol I to the Geneva Conventions is important for ensuring that a State's armed forces are capable of conducting hostilities in accordance with its international obligations," which is problematic with regard to AWS, firstly, because "the legal review must demand a very high level of confidence that, once activated, the autonomous weapon system would predictably and reliably operate as intedned,"[375] secondly, because software-based "AWS and remotely controlled weapon systems may appear identical from the outside" and

"cheating would be all too easy since the software could be changed back within minutes after inspection."[376]

Responsibility gap

"Individual and state responsibility is fundamental to ensure accountability for violations of international human rights and international humanitarian law."[377] It is necessary to hold war criminals accountable through the responsibility attribution, and the deployment of AWS will make that "so much harder"[378] because their use in combat is likely to create a "responsibility gap,"[379] or, alternatively – a "gap in accountability."[380]

The debate is whether designers, robot manufacturers, procurement officers, robot controllers or supervisors, field commanders, state governments or presidents, or even robots themselves should be held accountable.[381] There is a double-edge challenge because this is the first time "a weapon system will have either no one or too many people to be held accountable for mistakes."[382] There is a risk that military personnel may be held responsible for the actions of machines whose decisions and performance they barely control at best.[383] At the same time it is "hard to take seriously the idea that a machine should — or could — be held responsible" for the consequences of its own actions because "those who are punished, or contemplate punishment, should suffer as a result" where such suffering must be "morally compelling."[384] The attempts to punish a machine will also "have limited deterrent effects, since one robot could not be deterred by the punishment of another robot."[385]

This problem is further compounded by the "atomized approach of the law to questions of responsibility," which seeks to link a concrete and definable entity with some created specified effect, because it "runs contrary to the development of networks and swarms."[386] In any "system of systems," attempts to draw distinctions between the components, including between lethal and non-lethal systems, especially with regard to the allocation of criminal responsibility, will become increasingly arbitrary.[387]

The counter-argument calls to consider AWS in an unbiased manner with regard to the compliance with law and insists that, "as with most other weapon systems, their lawfulness as such, as well as the lawfulness of their use, must be judged on a case-by-case basis."[388]

Biased framing of AWS (non-)compliance with international law versus conventional warfare

The general argument is that it is true that the unlawful use of lawful weapons is not a rare phenomenon in the contemporary warfare.[389] With regard to IHL specifically, by itself, "autonomy is unlikely to present unnecessary suffering and superfluous injury issues since the rule addresses a weapon system's effect on the targeted individual, not the manner of engagement (autonomous)."[390] Moreover, the discrimination of combatants and non-combatants and the proportionality judgement are "no less difficult in air strikes and long-range attacks than they are with AWS."[391] The "fog of war" and "the lack of perfect situational awareness" dramatically complicate way to comply with the rules of international armed conflict for soldiers.[392] In turn, as it has been discussed in greater detail in the previous sections, it may be possible to program autonomous weapon systems in such a way that these machines through avoiding many of human mistakes and failings will potentially be able to "outperform human soldiers with respect to conformance to IHL."[393]

Human responsibility for AWS performance

The argument that there is no responsibility for AWS performance in the battlefield has "an air of triviality" because even "if the system is autonomous, it is not autonomous to the extent that it is completely independent of human authorship."[394] As any machine is programmed and deployed by human beings, the "responsibility for its operations lies unconditionally with them."[395] To make it clear, in AWS warfare there will always be "a human ultimately responsible for launching the weapon and putting it into operation, just not selecting the specific target."[396]

Concluding remarks

The presented article was aimed at restructuring and operationalising the debate on autonomous weapon systems, thereby allowing the reader for a balanced assessment of the issue on multiple key fronts. Our motivation was clear: within the broad realm of international security studies, there had not been such a detailed, balanced, and systematic analysis focused on the nitty-gritties of the AWS, their pros and cons being present side to side, and implications concerning their use being flagged. Because the issue of AWS has been heavily charged with highly

normative, moral, assertive and alarmist language from the very beginning, it was harder than in other security fields to get the facts right, and get to the level of genuine arguments to discuss characteristics of these weapons before pushing for a robust regulation, namely preventive, wide sweeping, globally working ban. However, the practice of fantasizing about the "killer robots" and spreading cultural myths and (mis)representational idioms of the Hollywood movies to create a media-rich spectre of danger has increasingly been counter-balanced with the reliance on technical knowledge coupled with a legal analysis and the workings of the security realm. This produced the debate flourishing at the two opposite poles, each being to a large degree aware of the counter-arguments put forward by the other side.

Unfortunately still, not all of the people involved in the so-called "killer robots" debate have actually had a balanced and detailed knowledge of the workings of these weapons, their precise delimitation, the novelty vs. continuity technology, and many other preconditions for a nuanced and informed opinion. By abandoning the one-sided terrain of an increasing number of NGOs, academics, politicians and scientists involved in the discussion and at the same making a step forward from what other neutrally analytical literature pieces had offered, we wished to create a first-order structure with a series of arguments and noteworthy counter-arguments along which the reader could easily navigate himself/herself. Our contribution to the debate, and a belief that it may help for its further cultivation and sophistication, was 1) to properly delimit the AWS category as there were too many flaws in this regard in the general debate – for us, that was to discuss fully autonomous (lethal) weapon systems, capable of operating without human control or supervision, including in dynamic and unstructured environments, and capable of engaging in independent (lethal) decision-making, targeting, and firing, including in an offensive manner; 2) to operationalise the issue of AWS along six dilemmas that we proposed as the basis, each of the dilemmas containing the detailed pro/con arguments – here, we went from the discussion of (un)predictability of AWS performance, dehumanization of lethal decision-making, depersonalisation of enemy (non-)combatant, human-machine nexus in coordinated operations, to strategic considerations and AWS operation in law(less) zone; 3) to bring on board literature from diverse fields to enrich intellectually what the plethora of media channels offer, namely robotics, computer science, law, security and strategic

studies, military ethics, philosophy and connection science. All of this has been done for the reader, in the first place, to be able to establish his/her own well-informed position on the issue through better understanding the nature of AWS and grasping the complexity of the debate. Potentially, this analytical framework, with each of its aspects potentially deserving its own place in future research, can also serve as the starting point for detecting and profoundly analysing multiple interlinkages between the dilemmas and (sub)sections in greater depth and in multiple possible ways. More generally, this article is intended to contribute to the better comprehension of AWS and more general challenges related to human-machine nexus and artificial intelligence.

Anzhelika Solovyeva is PhD Candidate at the Institute of Political Studies, Faculty of Social Sciences, Charles University in Prague and Doctoral Researcher at UNCE PERICULUM. She may be reached at anzhelika.solovyeva@fsv.cuni.cz

Nik Hynek is Associate Professor specialising in security studies and affiliated to Metropolitan University Prague and Charles University in Prague. Among others, he researches on Human-Machine Nexus and Its Implications for International Order. He may be reached at hynek@mup.cz and hynek@fsv.cuni.cz

The authors gratefully acknowledge funding for this work from Charles University in Prague, UNCE Human-Machine Nexus and Its Implications for International Order (UNCE/HUM/037).

Notes

1 Michael N. Schmitt (2013), 'Autonomous Weapon Systems and International Humanitarian Law: A Reply to the Critics,' *Harvard National Security Journal Feature*, p.2.
2 International Committee of the Red Cross (2014), *Autonomous Weapons Systems: Technical, Military, Legal, and Humanitarian Aspects*, report, p.7
3 US Department of Defense (2012), *Autonomy in Weapon Systems*, Directive 3000.09, Glossary, Part II Definitions, p.13.
4 Noel Sharkey (2017), 'Why Robots Should not be Delegated with the Decision to Kill,' *Connection Science* 29(2), pp.177-186, p.178.
5 Izumi Nakamitsu (2017), 'Foreword,' in *Perspectives on Lethal Autonomous Weapon Systems*, UNODA Occasional Papers No.30, New York: United Nations Publication, p.v.

6 Peter Asaro (2012), 'On Banning Autonomous Weapon Systems: Human Rights, Automation, and the Dehumanization of Lethal Decision-Making,' *International Review of the Red Cross* 94(886), pp.687-709, p.688.
7 Sharkey (2017), p.183.
8 Jürgen Altmann and Frank Sauer (2017), 'Autonomous Weapon Systems and Strategic Stability,' *Survival* 59(5), pp.117-142, p.132.
9 Human Rights Watch (2012), *Losing Humanity: The Case against Killer Robots*, report.
10 The International Committee on Robot Arms Control (ICRAC) was formed to promote the ban, see Jürgen Altmann, Peter Asaro, Noel Sharkey and Robert Sparrow (2009), *Mission Statement of the International Committee for Robot Arms Control*, available at: http://icrac.net/statements/; also see, for example, Human Rights Watch (2012); Asaro (2012), p.709; Noel E. Sharkey (2012), 'The Evitability of Autonomous Robot Warfare,' *International Review of the Red Cross* 94(886), pp.787-799, p.799; Aaron M. Johnson and Sidney Axinn (2013), 'The Morality of Autonomous Robots,' *Journal of Military Ethics* 12(2), pp.129-141, p.137; Human Rights Watch (2014), *Shaking the Foundations: The Human Rights Implications of Killer Robots*, report; 'Autonomous Weapons: an Open Letter from AI & Robotics Researchers' (2015), Future of Life Institute, available at: http://futureoflife.org/open-letter-autonomous-weapons/ [hereinafter: Open Letter]; Denise Garcia (2015), 'Killer Robots: Why the US should Lead the Ban,' *Global Policy* 6(1), pp.57-63, p.57; Human Rights Watch (2016), *Making the Case: The Dangers of Killer Robots and the Need for a Preemptive Ban*, report.
11 Ronald Arkin (2018), 'Lethal Autonomous Systems and the Plight of the Non-combatant,' in *The Political Economy of Robots*, edited by Ryan Kiggins, Switzerland: Palgrave Macmillan – International Political Economy Series, pp.317-326, p.321.
12 Kenneth Anderson and Matthew Waxman (2012), 'Law and Ethics for Robot Soldiers,' *Policy Review*, Issue 176, pp.35-49, pp.39-45.
13 Christof Heyns (2013), *Report of the Special Rapporteur on Extrajudicial, Summary or Arbitrary Executions*, United Nations Doc. A/HRC/23/47, p.5; Ian GR Shaw (2017), 'Robot Wars: US Empire and Geopolitics in the Robotic Age,' *Security Dialogue* 48(5), pp.451-470, pp.452-453; Arkin (2018), p.317.
14 Patrick Lin, George Bekey and Keith Abney (2008), *Autonomous Military Robotics: Risk, Ethics, and Design*, US Department of Navy, Office of Naval Research report, California Polytechnic State University: San Luis Obispo, p.4; Tim McFarland (2015), 'Factors Shaping the Legal Implications of Increasingly Autonomous Military Systems,' *International Review of the Red Cross* 97(900), pp.1313–1339, p.1324.
15 Lin, Bekey, Abney (2008), p.4.
16 Arkin (2018), p.317.
17 Sharkey (2012), p.797.
18 Noel Sharkey (2010), 'Saying 'No!' to Lethal Autonomous Targeting,' *Journal of Military Ethics* 9(4), pp.369-383, p.370.
19 Sharkey (2012), p.790.
20 Lin, Bekey, Abney (2008), p.87.
21 James I. Walsh (2015), 'Political Accountability and Autonomous Weapons,' *Research and Politics* 2(4), pp.1-6, p.2.

22 See, for example, the essays (2017) in *Perspectives on Lethal Autonomous Weapon Systems*, UNODA Occasional Papers No.30, New York: United Nations Publication.
23 See, for example, George R. Lucas (2010), 'Postmodern War,' *Journal of Military Ethics* 9(4), pp.289-298, p.293.
24 See, for example, Asaro (2012), p.693.
25 See, for example, Paul Scharre (2018), *Army of None: Autonomous Weapons and the Future of War*, W.W. Norton, Chapter 17.
26 See, for example, Sharkey (2012), p.790.
27 See, for example, Human Rights Watch (2012).
28 See, for example, Sharkey (2010), p.370.
29 See, for example, Mary Ellen O'Connell (2014), '21st Century Arms Control Challenges: Drones, Cyber Weapons, Killer Robots, and WMDs,' *Washington University Global Studies Law Review* 13(3), pp.515-533, p.526.
30 Altmann, Sauer (2017), p.132.
31 Ajey Lele (2017), 'A Military Perspective on Lethal Autonomous Weapon Systems,' in *Perspectives on Lethal Autonomous Weapon Systems*, UNODA Occasional Papers No.30, New York: United Nations Publication, pp.57-68, pp.58-59.
32 George R. Lucas (2014), 'Automated Warfare,' *Stanford Law and Policy Review*, Vol.25, pp.317-339, p.319.
33 Alex Leveringhaus (2016), *Ethics and Autonomous Weapons*, UK: Palgrave Macmillan, p.3
34 Lele (2017), p.59.
35 Heyns (2013), p.5.
36 ibid.
37 Asaro (2012), p.693.
38 Leveringhaus (2016), p.3.
39 Altmann, Sauer (2017), p.118.
40 McFarland (2015), p.1314.
41 Johnson, Axinn (2013), pp.137-138.
42 Altmann, Sauer (2017), p.118.
43 Lele (2017), p.59.
44 Guglielmo Tamburrini (2016), 'On Banning Autonomous Weapons Systems: From Deontological to Wide Consequentialist Reasons,' in Nehal Bhuta et al. (eds), *Autonomous Weapons Systems: Law, Ethics, Policy*, Cambridge: Cambridge University Press, pp.122-142, p.126.
45 Walsh (2015), p.2.
46 Altmann, Sauer (2017), p.118.
47 Benjamin Kastan (2013), 'Autonomous Weapons Systems: A Coming Legal 'Singularity'?,' *University of Illinois Journal of Law, Technology and Policy*, Issue 1, pp.45-82, p.49.
48 Altmann, Sauer (2017), p.118.
49 Asaro (2012), p.690.
50 Sharkey (2012), p.787.
51 Sharkey (2010), p.370.
52 Open Letter
53 See, for example, O'Connell (2014), p.526; Leveringhaus (2016), p.3; Michael C. Horowitz (2016), 'Why Words Matter: The Real World Consequences of Defining Autonomous Weapons Systems,' *Temple International and Comparative Law Journal* 30(1), pp.85-98, p.90; Denise Garcia (2018), 'Lethal

Artificial Intelligence and Change: The Future of International Peace and Security,' *International Studies Review* 20(2), pp.334-341, p.334.
54 Walsh (2015), p.2.
55 Tamburrini (2016), p.140.
56 Karolina Zawieska (2017), 'An Ethical Perspective on Autonomous Weapon Systems,' in *Perspectives on Lethal Autonomous Weapon Systems*, UNODA Occasional Papers No.30, New York: United Nations Publication, pp.49-56, p.49.
57 Asaro (2012), pp.708-709.
58 Tamburrini (2016), p.140.
59 Zawieska (2017), p.49.
60 Ronald C. Arkin (2017), 'A Roboticist's Perspective on Lethal Autonomous Weapon Systems,' in *Perspectives on Lethal Autonomous Weapon Systems*, UNODA Occasional Papers No.30, New York: United Nations Publication, pp.35-48, pp.38-39.
61 Arkin (2018), p.321.
62 Robert Sparrow (2007), 'Killer Robots,' *Journal of Applied Philosophy* 24(1), pp.62-77, p.65; McFarland (2015), p.1338;
63 Robert Sparrow (2009), 'Predators or Plowshares? Arms Control of Robotic Weapons,' *IEEE Technology and Society Magazine* 28(1), pp. 25-29, p.25.
64 Lin, Bekey, Abney (2008), p.78.
65 Michał Klincewicz (2015), 'Autonomous Weapons Systems, the Frame Problem and Computer Security,' *Journal of Military Ethics* 14(2), pp.162–176, p.167.
66 Sharkey (2012), p.789.
67 Paul Scharre (2011), 'Why unmanned,' *Joint Force Quarterly*, Issue 61, 2nd Quarter, pp.89-93, p. 92.
68 McFarland (2015), p.1335.
69 Lin, Bekey, Abney (2008), p.20.
70 Asaro (2012), p.691.
71 Klincewicz (2015), pp.165-166.
72 ibid, p.165.
73 Kareem Ayoub and Kenneth Payne (2016), 'Strategy in the Age of Artificial Intelligence,' *Journal of Strategic Studies* 39(5/6), pp.793-819, p.807.
74 Lin, Bekey, Abney (2008), p.78.
75 Jean Scholtz (2003), *Theory and Evaluation of Human Robot Interactions*, USA: IEEE Computer Society – Proceedings of the 36th Annual Hawaii International Conference on System Sciences.
76 Lin, Bekey, Abney (2008), p.40.
77 Ayoub, Payne (2016), p.793; Leveringhaus (2016), p.7
78 Armin Krishnan (2009), *Killer Robots: Legality and Ethicality of Autonomous Weapons*, Farnham: Ashgate Publishing, p.47; Human Rights Watch (2012); Lucas (2014), pp.327-328.
79 Mark Gubrud (2014), 'Stopping killer robots,' *Bulletin of the Atomic Scientists* 70(1), pp.32-42, p.33.
80 Krishnan (2009), p.47; Ayoub, Payne (2016) p.812; Leveringhaus (2016), p.7.
81 Ayoub, Payne (2016), pp.793-794.
82 A. Matthias (2004), 'The Responsibility Gap in Ascribing Responsibility for the Actions of Automata,' *Ethics and Information Technology* 6(3), pp.175–183, p.177.
83 Lin, Bekey, Abney (2008), p.20.

84 ibid, p.78.
85 ibid, p.40.
86 ibid, p.8.
87 Leveringhaus (2016), p.7.
88 Ayoub, Payne (2016) p.816.
89 Lin, Bekey, Abney (2008), p.66.
90 Peter Asaro (2008), 'How Just could a Robot War be?' in Adam Briggle, Katinka Waelbers and Philip A. E. Brey (eds), *Current Issues in Computing And Philosophy*, Amsterdam: IOS Press, pp. 50–64, p.55.
91 Gregory P. Noone and Diana C. Noone (2015), 'Debate Over Autonomous Weapons Systems,' *Case Western Reserve Journal of International Law* 47(1), pp.25-35, p.33.
92 Klincewicz (2015), p.168.
93 See, for example, Lin, Bekey, Abney (2008), p.78; Heyns (2013), p.18; Johnson, Axinn (2013), p.133; Noone, Noone (2015), p.33; McFarland (2015), p.1338; Dieter Birnbacher (2016), 'Are Autonomous Weapons Systems a Threat to Human Dignity?' in Nehal Bhuta et al. (eds), *Autonomous Weapons Systems: Law, Ethics, Policy*, Cambridge: Cambridge University Press, pp.105-121, p.118.
94 Lin, Bekey, Abney (2008), p.8.
95 ibid, p.79.
96 Klincewicz (2015), p.168.
97 Lin, Bekey, Abney (2008), p.79.
98 Noone, Noone (2015), p.33.
99 Klincewicz (2015), pp.169-170.
100 ibid, p.163.
101 Schmitt (2013), p.7.
102 Ayoub, Payne (2016), p.799.
103 Noel Sharkey (2008), 'Cassandra or False Prophet of Doom: AI Robots and War,' *IEEE Intelligent Systems* 23(4), p.16.
104 Gubrud (2014), p.33.
105 Paul Scharre (2017), 'A Security Perspective: Security Concerns and Possible Arms Control Approaches,' in *Perspectives on Lethal Autonomous Weapon Systems*, UNODA Occasional Papers No.30, New York: United Nations Publication, pp.19-34, p.26.
106 Sharkey (2017), p.182.
107 Lin, Bekey, Abney (2008), p.20.
108 Kastan (2013), p.53; Altmann, Sauer (2017), p.128.
109 Sharkey (2017), p.182.
110 ibid, p.183.
111 McFarland (2015), p.1323.
112 Schmitt (2013), p.7; McFarland (2015), p.1338.
113 Klincewicz (2015), p.172.
114 ibid, p.164.
115 Noone, Noone (2015), p.30.
116 Sparrow (2009), p.26.
117 McFarland (2015), pp.1326-1327.
118 Schmitt (2013), p.4.
119 McFarland (2015), pp.1327-1328.
120 ibid, p.1335.
121 Michael Robillard (2017), 'No Such Thing as Killer Robots,' *Journal of Applied Philosophy*, p.7.

122 Kastan (2013), p.50.
123 Lucas (2014), pp.323-324.
124 Markus Wagner (2014), 'The Dehumanization of International Humanitarian Law: Legal, Ethical, and Political Implications of Autonomous Weapon Systems,' *Vanderbilt Journal of Transnational Law* 47(5), pp.1371-1424, p.1410.
125 Sharkey (2012), p.799; Scharre (2018), Chapter 17.
126 Srđan T. Korać (2018), 'Depersonalisation of Killing: Towards A 21st Century Use Of Force ‹Beyond Good And Evil?',' *Philosophy and Society* 29(1), pp.49-64, p.62.
127 Ayoub, Payne (2016) p.798.
128 Wagner (2014), p.1415.
129 Human Rights Watch, (2012); Johnson, Axinn (2013), p.136.
130 See, for example, Krishnan (2009), p.130; Lucas (2014), p.327; Wagner (2014), pp.1410 and 1415; Birnbacher (2016), p.120; Human Rights Watch (2016); Scharre (2018), Chapter 17; Korać (2018), p.51.
131 Korać (2018), p.51.
132 Human Rights Watch (2012); Human Rights Watch (2016).
133 Johnson, Axinn (2013), p.129; Ayoub, Payne (2016) p.807.
134 Scharre (2018), Chapter 17.
135 Human Rights Watch (2016).
136 Birnbacher (2016), p.121 (Referring to Christof Heyns, the Florence conference).
137 Lin, Bekey, Abney (2008), p.32.
138 Asaro (2012), p.701; Sharkey (2017), p.180.
139 Asaro (2008), p.61.
140 ibid, p.60.
141 Kastan (2013), p.60.
142 Noel Sharkey (2008a), 'Grounds for Discrimination: Autonomous Robot Weapons', *RUSI Defence Systems* 11(2), p.86.
143 Asaro (2008), p.60.
144 Kastan (2013), p.60.
145 ibid, p.48.
146 Noone, Noone (2015), p.26.
147 Asaro (2012), p.700.
148 ibid.
149 Asaro (2008), p.60.
150 Johnson, Axinn (2013), p.135.
151 Johnson, Axinn (2013), p.135; Robillard (2017), p.6.
152 Asaro (2012), p.701.
153 Heyns (2013), p.14.
154 Asaro (2012), p.688.
155 Robert Sparrow (2009a), 'Building a Better WarBot: Ethical Issues in the Design of Unmanned Systems for Military Applications,' *Science and Engineering Ethics* 15(2), pp.169-187, p.178; Hin-Yan Liu (2012), 'Categorization and Legality of Autonomous and Remote Weapons Systems,' *International Review of the Red Cross* 94(886), p.645; Johnson, Axinn (2013), p.137; Korać (2018), p.59.
156 Lin, Bekey, Abney (2008), p.65; Johnson, Axinn (2013), p.135.
157 Ryan Tonkens (2012), 'The Case against Robotic Warfare: A Response to Arkin,' *Journal of Military Ethics* 11(2), pp.149-168, p.151.

158 Johnson, Axinn (2013), pp.135-136.
159 ibid, p.129.
160 Birnbacher (2016), p.121.
161 ibid.
162 Ronald Arkin (2010), 'The Case for Ethical Autonomy in Unmanned Systems,' *Journal of Military Ethics* 9(4), pp. 332–341, p.334.
163 Arkin (2018), p.318
164 Lin, Bekey, Abney (2008), p.1.
165 See, primarily, Gubrud (2014), p.38; see also, for example, Lin, Bekey, Abney (2008), p.1; Arkin (2010), p.337; Klincewicz (2015), p.164; Ayoub, Payne (2016) p.798.
166 Schmitt (2013), p.13.
167 Klincewicz (2015), p.164.
168 Birnbacher (2016), p.119.
169 See, for example, Arkin (2010), p.333; Heyns (2013), p.10; Lucas (2014), p.326; Noone, Noone (2015), pp.29-30; Birnbacher (2016), p.119; Arkin (2018), p.318.
170 Arkin (2010), p.337.
171 Arkin (2010), p.333; Arkin (2018), p.320.
172 Arkin (2010), p.333; Noone, Noone, (2015), p.32; Korać (2018), pp.56.
173 Birnbacher (2016), p.119.
174 See, for example, Scholtz (2003); Lin, Bekey, Abney (2008), p.1; Arkin (2010), pp.333 and 336; Scharre (2011), p.92; Ayoub, Payne (2016), p.798; Sharkey (2017), p.183; Arkin (2018), pp.318-320.
175 Arkin (2010), p.338; Wagner (2014), p.1411; Noone, Noone (2015), p.31.
176 Noone, Noone (2015), p.32.
177 Scharre (2018), Chapter 17.
178 See, for example, Lucas (2014), p.326; Klincewicz (2015), p.164; Arkin (2018), p.319.
179 Altmann, Sauer (2017), p.119.
180 Liu (2012), p.649.
181 Lin, Bekey, Abney (2008), p.1.
182 Sparrow (2009), p.26.
183 Altmann, Sauer (2017), p.119.
184 Noone, Noone (2015), p.25.
185 Scharre (2018), Chapter 17.
186 Sharkey (2012), p.791
187 Ayoub, Payne (2016), p.799.
188 Scharre (2018), Chapter 17.
189 Lucas (2010), p.293.
190 Klincewicz (2015), p.164.
191 Noone, Noone (2015), p.29.
192 Arkin (2010), pp.333-334; Arkin (2018), p.321.
193 Tonkens (2012), p.151; Johnson, Axinn (2013), pp.135-136.
194 Noone, Noone (2015), p.33.
195 ibid.
196 Klincewicz (2015), p.163.
197 Christof Heyns (2016), 'Autonomous Weapons Systems: Living a Dignified Life and Dying a Dignified Death,' in Nehal Bhuta et al. (eds), *Autonomous Weapons Systems: Law, Ethics, Policy*, Cambridge: Cambridge University Press, pp. 3–20, p.12.

198 Korać (2018), p.54.
199 ibid, pp.49-60.
200 Sharkey (2012), p.788.
201 Korać (2018), p.58.
202 Heyns (2016), p.3.
203 Sparrow (2009), p.26; U.C. Jha (2016), *Killer Robots: Lethal Autonomous Weapon Systems: Legal, Ethical and Moral Challenges*, India: Vij Books India, Chapter 5.
204 Heyns (2016), p.4.
205 Garcia (2015), p.61.
206 Wagner (2014), p.1411.
207 Korać (2018), p.50.
208 ibid, p.59.
209 Sharkey (2010), p.369.
210 Sparrow (2009), p.26.
211 Scharre (2018), Chapter 17.
212 Sparrow (2009a), p.183.
213 Heyns (2016), p.11; Human Rights Watch (2016); Scharre (2018), Chapter 17.
214 Korać (2018), p.59.
215 Heyns (2016), pp.10-11.
216 McFarland (2015), p.1339.
217 ibid, p.1335.
218 Asaro (2008), p.60.
219 Sharkey (2012), p.789.
220 Birnbacher (2016), p.118.
221 See, for example, Sparrow (2007), p.68; Heyns (2013), p.6; Johnson, Axinn (2013), p.134; Human Rights Watch (2014); Garcia (2015), pp.58-61; Heyns (2016), p.3.
222 Birnbacher (2016), p.121.
223 Wagner (2014), p.1420.
224 Birnbacher (2016), p.119.
225 Scharre (2018), Chapter 17.
226 Noone, Noone (2015), p.33.
227 Birnbacher (2016), pp.118-121.
228 Noone, Noone (2015), p.33.
229 Birnbacher (2016), p.120.
230 Scharre (2018), Chapter 17.
231 Wagner (2014), p.1411.
232 Scharre (2018), Chapter 17.
233 Birnbacher (2016), p.119.
234 Heyns (2016), p.7.
235 Heyns (2013), p.9.
236 Schmitt (2013), p.6.
237 Wagner (2014), p.1419.
238 Sparrow (2009a), p.172.
239 Sparrow (2009), p.26.
240 Asaro (2008), p.62.
241 Scharre (2018), Chapter 17.
242 Sparrow (2009a), p.173.
243 ibid.
244 ibid, pp.172-173.

245 Lin, Bekey, Abney (2008), p.80.
246 Liu (2012), p.633.
247 Sparrow (2009a), p.173.
248 Sharkey (2012), p.792.
249 Sparrow (2007), p.64.
250 Lin, Bekey, Abney (2008), p.20.
251 Asaro (2008), p.61.
252 Kastan (2013), p.53.
253 Johnson, Axinn (2013), p.134.
254 Ayoub, Payne (2016), p.797.
255 Lin, Bekey, Abney (2008), p.74.
256 ibid, pp.79-80.
257 Heyns (2016), p.6.
258 Arkin (2017), p.37.
259 Arkin (2018), p.318
260 Johnson, Axinn (2013), p.133.
261 Noone, Noone (2015), p.33.
262 Arkin (2010), pp.336-337.
263 Noone, Noone (2015), p.32.
264 Kastan (2013), p.54; Heyns (2016), p.7; Altmann, Sauer (2017), p.125.
265 Klincewicz (2015), p.164.
266 Ayoub, Payne (2016), p.807.
267 ibid, p.816.
268 Klincewicz (2015), p.164.
269 Ayoub, Payne (2016), p.793.
270 Leveringhaus (2016), p.7.
271 Scharre (2017), p.22.
272 Heather M. Roff (2014), 'The Strategic Robot Problem: Lethal Autonomous Weapons in War,' *Journal of Military Ethics* 13(3), pp.211-227, p.212.
273 Korać (2018), pp.57-58.
274 Sharkey (2017), p.183.
275 Gubrud (2014), p.39.
276 Ayoub, Payne (2016), p.814.
277 ibid, pp.796-797.
278 ibid, p.797.
279 Sharkey (2008), p.16.
280 Garcia (2018), p.335.
281 Asaro (2012), p.692.
282 Liu (2012), p.648.
283 Sparrow (2007), p.69.
284 Gubrud (2014), p.39.
285 Altmann, Sauer (2017), p.132.
286 Open Letter
287 Altmann, Sauer (2017), p.127.
288 Heyns (2013), p.16.
289 ibid.
290 Korać (2018), p.61.
291 See, for example, Heyns (2013), p.18; Open Letter; Altmann, Sauer (2017), p.127.
292 J.S. Canning (2002), *A Definitive Work on Factors Impacting the Arming of Unmanned Vehicles*, report NSWCDD/TR-05/36 P, Dahlgren Division, Naval Surface Warfare Center, p.32.

293 Wagner (2014), p.1420.
294 ibid, p.1419.
295 Gubrud (2014), p.39.
296 Korać (2018), p.54.
297 Asaro (2012), p.692; Heyns (2013), p.11; Sharkey (2017), p.182.
298 Heyns (2013), p.11.
299 ibid.
300 ibid.
301 Korać (2018), p.61.
302 Sparrow (2009), p.27.
303 Sharkey (2017), p.183.
304 Sparrow (2009), p.26.
305 Gubrud (2014), p.39.
306 Sparrow (2009), p.27.
307 Asaro (2012), p.692.
308 Liu (2012), p.648.
309 Asaro (2012), p.693.
310 Heyns (2013), p.13.
311 Asaro (2008), p.63.
312 Garcia (2018), p.339.
313 Asaro (2008), p.62.
314 Heyns (2013), p.16.
315 ibid, p.12.
316 ibid, p.16.
317 Sharkey (2008), p.16; Heyns (2013), p.16.
318 Asaro (2008), p.63.
319 Lin, Bekey, Abney (2008), p.81.
320 Schmitt (2013), p.7.
321 Lin, Bekey, Abney (2008), p.80.
322 Kastan (2013), p.53.
323 Ayoub, Payne (2016), p.816.
324 Scharre (2017), p.22.
325 Liu (2012), p.633; Altmann, Sauer (2017), p.119.
326 Ayoub, Payne (2016), pp.807-808.
327 Lin, Bekey, Abney (2008), p.1; Heyns (2013), p.10; Arkin (2017), p.36; Arkin (2018), p.318.
328 Lin, Bekey, Abney (2008), p.1.
329 Heyns (2013), p.10.
330 Sharkey (2008a), p.87.
331 Liu (2012), p.633.
332 Wagner (2014), p.1413.
333 ibid.
334 Ayoub, Payne (2016), p.807.
335 Korać (2018), p.57.
336 Kastan (2013), p.52; Lucas (2014), p.325; Arkin (2018), p.318.
337 Ayoub, Payne (2016), p.808.
338 Arkin (2017), p.37.
339 Heyns (2013), p.10.
340 Sparrow (2007), p.68.
341 Altmann, Sauer (2017), p.119.
342 Sparrow (2009), p.27.

343 Gubrud (2014), p.38.
344 Wagner (2014), p.1420.
345 Liu (2012), p.633.
346 Lucas (2014), p.326.
347 Lin, Bekey, Abney (2008), p.54.
348 Wagner (2014), p.1411.
349 ibid.
350 Leveringhaus (2016), p.7.
351 Ayoub, Payne (2016), p.797.
352 Leveringhaus (2016), p.7.
353 Asaro (2008), pp.62-63.
354 Heyns (2013), p.12.
355 Lin, Bekey, Abney (2008), p.54.
356 Kastan (2013), p.47.
357 Lin, Bekey, Abney (2008), p.88.
358 ibid.
359 Wagner (2014), p.1371.
360 Heyns (2013), p.11.
361 Garcia (2015), p.60.
362 ibid.
363 Human Rights Watch (2016).
364 Human Rights Watch (2014); Garcia (2015), pp.60-61.
365 Heyns (2016), pp.8-10.
366 ibid.
367 McFarland (2015), p.1337.
368 Korać (2018), p.57.
369 Wagner (2014), pp.1384-1385.
370 Kastan (2013), p.60.
371 Lin, Bekey, Abney (2008), p.76.
372 Heyns (2013), p.13.
373 ibid, p.11.
374 Walsh (2015), p.5.
375 Neil Davison (2017), 'A Legal Perspective: Autonomous Weapon Systems under International Humanitarian Law,' in *Perspectives on Lethal Autonomous Weapon Systems*, UNODA Occasional Papers No.30, New York: United Nations Publication, pp.5-18, pp.9-10.
376 Altmann, Sauer (2017), p.135.
377 Heyns (2013), p.14.
378 Garcia (2018), p.339.
379 Matthias (2004), p.177; Liu (2012), p.630.
380 Human Rights Watch (2016).
381 Lin, Bekey, Abney (2008), p.73.
382 Garcia (2015), p.60.
383 Sparrow (2007), p.71.
384 ibid, pp.71-72.
385 Kastan (2013), p.68.
386 Liu (2012), p.650.
387 McFarland (2015), p.1333.
388 Schmitt (2013), p.8.
389 ibid, p.14.
390 ibid, p.9.

391 Birnbacher (2016), p.119.
392 Korać (2018), p.56.
393 Arkin (2018), p.323.
394 Birnbacher (2016), p.120.
395 ibid.
396 Scharre (2018), Chapter 17.