

Autonomous Weapon Systems, *Quo Vadis?*

Iness Arabi

IE University, Madrid, Spain

E-mail: iarabi.ieu2016@student.ie.edu

Published: April 10, 2019



Abstract

This paper examines the implications arising from states' use of autonomous weapon systems in situations of conflict. The analysis starts by addressing the definitional problems found in the literature on autonomous weapon systems. The primary finding is that the differential feature of autonomous weapon systems is their ability to select among targets and decide to kill without human oversight. The paper then delves into the effect that the increasing use of autonomous weapon systems has on conflict and war and the resulting policy implications for states and the international community as a whole. I conclude by discussing the legal, ethical, and moral implications of the use of weapons that can kill autonomously, which are at the core of the debate.

Keywords: autonomous weapon systems, unmanned weapons, foreign policy, law of armed conflicts.

1. Introduction

We find ourselves in a time in which the rapid advances of technology profoundly affect, if not completely revolutionize, how the world operates. From the rise of cyberterrorism to the effects of social media on democracy, international relations have not remained unscathed. Even so, the most Terminator-like concern has been one raised in the last decade: the creation of *killer robots*¹. What could have well been the plot of a science-fiction movie is now the concern of academics and policy-makers alike.

The revolutionary effect of autonomous weapons systems on warfare and state relations has been likened to that of gunpowder, computers, and even electricity.² In the face of such sensationalism, we must ask ourselves: why are autonomous weapons set to change the world as we know it? More importantly, how do we ensure that we are two steps ahead of these *killer robots*?

This paper will answer these questions by addressing the following issues. First, in an aim to bring clarity to what autonomous weapon systems are, I will address the definitional problems found in the literature on autonomous weapon systems. More specifically, I will analyze the notion of "autonomy" and where different stakeholders draw the line of autonomy. Second, this paper will address the policy implications of autonomous weapon systems. Finally, I will address the ethical, legal and moral concerns

raised in the public debate on autonomous weapon systems.

2. Autonomous Weapon Systems: What is in a name?

The advent of autonomous weapon systems has been given much momentum in public policy and has been closely anticipated and monitored. Part of the reason for this is the widespread belief that these *killer robots* are unique and revolutionary. If this is the case, we must ask ourselves why that is. What exactly makes these weapons so different from their predecessors? It would seem that the answer lies in their "autonomy."

1. Existing definitions

The United States Department of Defense has defined autonomous weapon systems as systems that *"once activated, can select and engage targets without further intervention by a human operator. This includes human-supervised autonomous weapon systems that are designed to allow human operators to override operation of the weapon system, but can select and engage targets without further human input after activation."*³

Conversely, semi-autonomous weapon systems as systems that *"once activated, [are] intended to only engage individual targets or specific target groups that have been selected by a human operator."*⁴ The main point is that *"human control is retained over the decision to select individual*

¹ Crootof, Rebecca. "The killer robots are here: legal and policy implications." *Cardozo L. Rev.* 36 (2014): 1837.

² Naval Research Committee: Autonomous and Unmanned Systems in the Department of the Navy

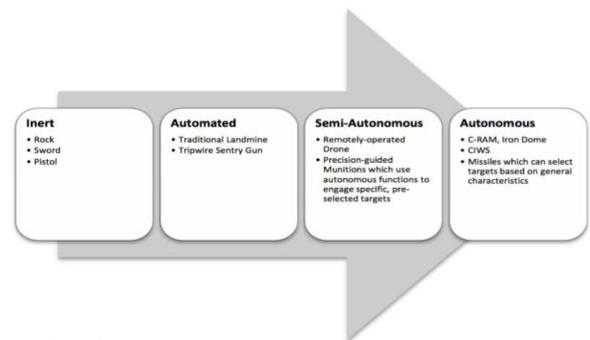
³ D United States, Department of Defense, Executive Service Directorate. "Department of Defense Directive 3000.09" Department of Defense Directive , ser. 3000.09, 2012.

⁴ *Ibid.*

*targets and specific target groups for engagement.*⁵

Some scholars⁶ have rightly pointed out that, in an abstract sense, weapons such as landmines could qualify as autonomous weapon systems under that definition, as they are triggered without a human operator. In other words, there is no human oversight over who the target is. Given this ambiguity, it has been necessary to narrow the function of “select” to “select among” targets. Under this development, “selection among” would entail that there is “a machine-generated targeting decision made; some form of computational cognition, meaning some form of AI or logical reasoning, is inherently part of autonomous weapon systems in the contemporary debate.”⁷ Consequently, autonomous weapon systems would possess “some decisional capability to ‘select’ and ‘engage.’”

Figure 1⁸



2. Drawing at the line at “autonomy”

According to these definitions, it would seem that the line of “autonomy” is drawn at the decision-making level and more specifically at the selection of targets. This distinction has been corroborated by the International Committee of the Red Cross, which has defined autonomous weapons systems as “any weapon system with autonomy in its critical functions—that is, a weapon system that can select (search for, detect, identify, track or select) and attack (use force against, neutralize, damage or destroy) targets without human intervention.”⁹

Alternatively, some authors¹⁰ have argued that a dichotomous division is not reflective of the practical reality of these weapons. Instead, the level of autonomy of different weapon systems will depend on the interactions between human operators and machine functions and should be

⁵ *Ibid.*

⁶ Anderson, Kenneth, and Matthew C. Waxman. "Debating Autonomous Weapon Systems, their Ethics, and their Regulation under international law." (2017).

⁷ Anderson, Kenneth, and Matthew C. Waxman. "Debating Autonomous Weapon Systems, their Ethics, and their Regulation under international law." (2017).

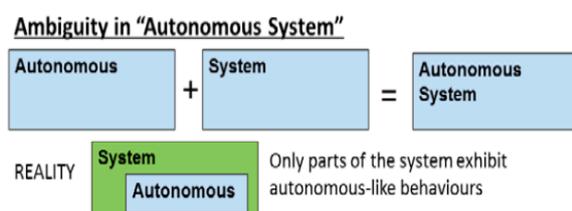
⁸ Figure 1: Crootof, Rebecca. "The killer robots are here: legal and policy implications." *Cardozo L. Rev.* 36 (2014): 1837.

⁹ Davison, Neil. "A legal perspective: Autonomous weapon systems under international humanitarian law." *Perspectives on lethal autonomous weapon systems* (2017): 5-18.

¹⁰ Anderson, Kenneth, and Matthew C. Waxman. "Debating Autonomous Weapon Systems, their Ethics, and their Regulation under international law." (2017).

assessed on a case-by-case basis. Others¹¹ have posited that the term “autonomous systems” creates confusion and ambiguity, as it clusters together systems that are fundamentally different by using “autonomy” as their main label over and above all of other features and capabilities.¹² These scholars have proposed to use an alternative nomenclature for these systems: “autonomous function in a system”.

Figure 2¹³



While there may be divergence in the literature regarding the definition and the nomenclature given to autonomous weapon systems, consensus can be found on the fact that ahead of us lie increases in levels of autonomy and decreases in levels of human intervention until the human role is negligibly small. In all likelihood, human intervention will be limited to activating the weapons.¹⁴

3. Policy Implications

One of the concerns raised by political scientists and policymakers is how the advent of autonomous weapon systems will impact the likelihood of conflict and war. The main argument here is that the development and use of lethal weapons that “pose little risk to the lives of the operators removes a potent deterrent for armed conflict”¹⁵ and will consequently “revolutionize warfare.”¹⁶ This revolution would come, on the one hand, from the decrease in the operational cost of war and would therefore “democratize” warfare by increasing the military capabilities of smaller states¹⁷ and, on the other hand, from the disappearance of the transaction cost that comes with sending troops to combat. The latter effectively de-politicizes the question of whether to go to war, as it stops being a high-cost issue for the constituency or a polarizing issue in public opinion. In other words, the concern is: what will warfare look like once it no longer is an issue of public debate?

The first implication, namely that of the “democratization” of warfare, could have profound implications for the global balance of power, similar but not to the extent of that of nuclear weapons. Additionally, many policy papers^{18,19} have warned against the effects that these weapons would have on global terrorism.

¹¹ Williams, Andrew. “Defining Autonomy in Systems: Challenges and Solutions.” *Issues for Defence Policymakers*(2015): 27.

¹² *Ibid.*

¹³ Figure 2.3: Williams, Andrew. “Defining Autonomy in Systems: Challenges and Solutions.” *Issues for Defence Policymakers*(2015): 27.

¹⁴ United States, Department of Defense, Executive Service Directorate. “Department of Defense Directive 3000.09” Department of Defense Directive , ser. 3000.09, 2012.

¹⁵ Scott, Ben, Stefan Heumann, and Philippe Lorenz. “Artificial Intelligence and Foreign Policy.” *Stiftung Neue Verantwortung Policy Brief* (2018).

¹⁶ *Ibid.*

¹⁷ Artificial Intelligence and National Security Greg Allen Taniel Chan A study on behalf of Dr. Jason Matheny, Director of the U.S. Intelligence Advanced Research Projects Activity (IARPA)

¹⁸ *Ibid.*

¹⁹ Williams, Andrew. “Defining Autonomy in Systems: Challenges and Solutions.” *Issues for Defence Policymakers*(2015): 27.

Some even contend that “a new arms race appears inevitable alongside a new set of dangers from terrorism.”²⁰

In the face of such possibilities, many have called for a complete ban of autonomous weapon systems. In fact, in 2015, an open letter signed by over three thousand leading AI researchers was presented at the International Joint Conference on Artificial Intelligence in Buenos Aires, Argentina, calling for a ban on offensive autonomous weapons. Other experts have taken more strategic approaches²¹ and have set out strategy plans to ensure their state’s superiority in the field. Others prefer a more *laissez-faire* approach by claiming that, because autonomous are already being used lawfully today, international law already regulates their creation, development, and use.

When it comes to ensuring a successful ban, Rebecca Crootof, expert on autonomous weapon systems and author of “Killer Robots”, has identified the different factors that have led to the ban of previous weapons and contends that at least one of these need to apply to ensure the practical and successful ban of any type of weapon system²²: weapons causing superfluous injury or unnecessary suffering, inherently indiscriminate weapons, ineffective weapons, other existing means for accomplishing the same military objective, clear and narrowly tailored prohibitions, prior regulation, public concern and civil society engagement, and sufficient state

²⁰ Scott, Ben, Stefan Heumann, and Philippe Lorenz. "Artificial Intelligence and Foreign Policy." *Stiftung Neue Verantwortung Policy Brief* (2018).

²¹ Artificial Intelligence and National Security Greg Allen Taniel Chan A study on behalf of Dr. Jason Matheny, Director of the U.S. Intelligence Advanced Research Projects Activity (IARPA)

²² Crootof, Rebecca. "The killer robots are here: legal and policy implications." *Cardozo L. Rev.* 36 (2014): 1837.

commitment. Crootof claims that the only factor applicable to the ban of autonomous weapon systems is “public concern and civil society engagement”, particularly because: (i) states already use autonomous weapon systems, and (ii) the most common concerns (which will be addressed later in this paper) are framed in ethical, legal or moral terms. Crootof draws the parallel with the Mine Ban Convention and the Convention on Cluster Munitions which has been attributed mainly to the participation of nongovernmental organizations and other civil society representatives.²³

4. Other considerations

The debate about autonomous weapon systems, which has spread to the realm of public opinion, has been framed in ethical, legal, and moral terms. Is it ethical for us to allow machines to decide on whom to target?²⁴ Are autonomous weapon systems in breach of the distinction principle of international humanitarian law?²⁵ These are the questions that one can find in the literature on autonomous weapon systems. This paper will continue by addressing the implications arising from such concerns.

1. Ethical, legal, and moral considerations

Regarding the ethical, legal, and political dilemmas that autonomous weapon systems seem to pose for a number of scholars, this paper will address the following ones: (i) do autonomous weapon systems currently fulfill the requirements of the law of armed conflicts in international

²³ *Ibid.*

²⁴ Crootof, Rebecca. "The killer robots are here: legal and policy implications." *Cardozo L. Rev.* 36 (2014): 1837.

²⁵ Crootof, Rebecca. "The killer robots are here: legal and policy implications." *Cardozo L. Rev.* 36 (2014): 1837.

humanitarian law to be lawfully used, and if not, will they ever?, (ii) do autonomous weapon systems hinder or impede accountability in armed conflicts?, and most importantly, (iii) do human beings have the moral monopoly on killing?

Many scholars^{26,27} have addressed the common and popular claim that autonomous weapon systems will never be able to comply with the law of armed conflict. I will proceed by deconstructing this claim.

First and foremost, it would seem that it rests on assumptions about how technology, artificial intelligence and weaponry will evolve in the future, and that is in a way that will never fulfill the set of requirements imposed by international humanitarian law. It is true that machines and weapon systems may never develop moral and ethical values. However, this should not give way to skeptical and unfounded assumptions on technological evolution. Instead, it should incentivize engineers, policy makers, and legal authorities alike to develop ways to circumvent this issue.

Second, it rests on assumptions on how international humanitarian law will evolve and, specifically, on its lack of flexibility. While it is true that many of the principles that are the backbone of international humanitarian law today have been in use for decades, if not centuries, the law has also proven to be flexible enough to address the emerging issues it has been faced with with time. If the law remains static while reality is in constant motion and evolution, we will find ourselves operating within an obsolete and

²⁶ Crootof, Rebecca. "The killer robots are here: legal and policy implications." *Cardozo L. Rev.* 36 (2014): 1837.

²⁷ Anderson, Kenneth, and Matthew C. Waxman. "Debating Autonomous Weapon Systems, their Ethics, and their Regulation under international law." (2017).

outdated framework. Moreover, an interesting reality that has been pointed out is that autonomous weapon systems are currently being employed lawfully, proving that they are not inherently unlawful.²⁸

Within this broader legal debate, much attention has been paid to the principle of distinction, namely the legal precept that differentiates between "military objectives and civilian objects, combatants and civilians, and active combatants and those hors de combat."²⁹ Military commanders and actors in conflict must abide by this principle, and by extension, so must autonomous weapon systems. On the one hand, most scholars and experts agree that autonomous weapon systems are incapable of distinguishing between combatants and civilians³⁰, thus rendering them unlawful under the distinction principle. On the other hand, some have raised doubts about the ability of humans to make such distinctions, especially in the fog of war. The difference, it would seem, between an autonomous legal system and a human commander, both of which do not abide by the distinction principle is that the human commander can be held accountable for a breach of international humanitarian law, while a machine cannot.

This takes us to our second concern, and that is whether the use of autonomous weapon systems could hinder accountability in the realm of armed conflict. The International Committee of the Red Cross has been very categorical in its view on this

²⁸ Crootof, Rebecca. "The killer robots are here: legal and policy implications." *Cardozo L. Rev.* 36 (2014): 1837.

²⁹ Davison, Neil. "A legal perspective: Autonomous weapon systems under international humanitarian law." *Perspectives on lethal autonomous weapon systems* (2017): 5-18.

³⁰ Crootof, Rebecca. "The killer robots are here: legal and policy implications." *Cardozo L. Rev.* 36 (2014): 1837.

issue and has stated that “all obligations under international law and accountability for them cannot be transferred to a machine, computer program or weapon system.”³¹ Consequently, these weapons “should be banned because machine decision-making undermines, or even removes, the possibility of holding anyone accountable in the way and to the extent that, for example, an individual human soldier might be held accountable for unlawful or even criminal actions.”³² This argument relies on the weight that individual criminal responsibility has on international law. While the importance of the emergence of individual criminal responsibility in the last half-century and the impact and contribution of its institutions (the International Criminal Court, the Nuremberg trials, the International Military Tribunal for the Far East, the International Criminal Tribunal for the former Yugoslavia, the International Criminal Tribunal for Rwanda, etc) are undeniable, it is also true that the “effective adherence to the law of armed conflict traditionally has come about through mechanisms of state (or armed party) responsibility.”³³ Thus, the use of autonomous weapon systems would not impede the establishment of criminal responsibility for the party that has unlawfully deployed them.

The last and perhaps most important question is, put in simple terms, whether machines can morally decide to kill. This question is vested on the underlying premise that human beings have

the monopoly on morality, and by extension, moral killing. Human beings have decided what is moral throughout time and space. More recently, social psychology has introduced the idea of “framing” as the way in which public opinion, and by extension, common notions of morality and ethics are framed. Consequently, we find that the initial question becomes obsolete, it is no longer an issue of whether machines are morally able to kill, but instead, whether machines can kill within the framework of morality created by human beings at a certain point in time and space.

The American roboticist Ronald C. Arkin has addressed this issue by developing the eponymous Arkin test, under which “an unmanned platform fulfills the demands of law and morality (and may therefore be permissibly employed) when it can be shown to comply with legal and moral requirements and constraints as well or better than a human under similar circumstances”³⁴. It seems that, nowadays, no machine passes the Arkin test. Currently, the largest effort to reproduce human conscience in a machine is in “strong AI”, which would replicate human decision-making processes and capabilities in machines. This raises the question, is this a good thing?

The premise behind the arguments in favor of “strong AI” and the Arkin test is that because human beings *can* act morally, they *do* act morally. Furthermore, it harbours the idea that these human capabilities somehow render decisions safer or more reliable, thus completely removing human failings and error out of the equation. This assumption ignores the flip side of the coin, which is that any notion of morality inherently carries with it notions of immorality. In other words, if human beings can be moral, they can also be immoral and act immorally. Machines,

³¹ Davison, Neil. "A legal perspective: Autonomous weapon systems under international humanitarian law." *Perspectives on lethal autonomous weapon systems* (2017): 5-18.

³² Anderson, Kenneth, and Matthew C. Waxman. "Debating Autonomous Weapon Systems, their Ethics, and their Regulation under international law." (2017).

³³ Anderson, Kenneth, and Matthew C. Waxman. "Debating Autonomous Weapon Systems, their Ethics, and their Regulation under international law." (2017).

³⁴ Lucas Jr, George R. "Automated Warfare." *Stan. L. & Pol'y Rev.* 25 (2014): 317

on the other hand, act and operate outside of the framework of morality. They, like animals, are *amoral*. So far, the amorality of machines has been implicitly equated to the immorality of humans, but these are profoundly distinct. As some scholars have pointed out, the fact that machines do not pass the Arkin test and may never pass the Arkin test gives us the reassurance that unmanned systems could not emulate any undesirable human reactions³⁵, which until now have been behind all military catastrophes. This is because machines “do not care, they have no interests, intentions, or self-regard, they harbor no ambitions or hatred, and they are utterly incapable of the “interiority” characteristic of self-consciousness.”³⁶ And so, we reach the conclusion that not only is it impossible for robots to be human, but that neither would we wish them to be.

5. Conclusion

Autonomous weapon systems have made headlines in the last decades causing equal amounts of outrage and praise among civil society and in academic debate. This is mainly due to their differential feature: autonomy. *Mutatis mutandis*, autonomous weapon systems have the ability to select among targets and decide to kill without any human intervention or oversight.

The arguments against the use of autonomous weapon systems are political, legal, and moral. Politically, it would seem that these weapons may incentivize states and non-state actors, such as terrorist groups, to turn to armed conflict. Legally, the lack of human oversight over decisive actions in conflict may impede the establishment of individual criminal responsibility. Morally, it would

seem that giving machines the power to decide on the life of a human being is wrong.

Proponents, on the other hand, refute these arguments and find that autonomous weapon systems may make conflict less costly and more efficient. Politically, the deployment of troops and the loss of casualties is reduced or even eliminated. Legally, the use of autonomous weapon systems does not affect the establishment of criminal responsibility of each party in armed conflict. Finally, morally, autonomous weapon systems substitute human emotions and interests by algorithms and lines of code, thus eliminating human error from the decision to kill. Outside of this debate, the reality is that states currently deploy autonomous weapon systems in combat. Civil society, however, remains strongly against their use and calls for a complete ban of these weapons. Only time will tell whether the people’s voices will be loud enough to be heard.

³⁵ *Ibid.*

³⁶ *Ibid.*

Bibliography

Crootof, Rebecca. "The killer robots are here: legal and policy implications." *Cardozo L. Rev.* 36 (2014): 1837.

United States, Department of Defense, Executive Service Directorate. "Department of Defense Directive 3000.09" *Department of Defense Directive*, ser. 3000.09, 2012.

Anderson, Kenneth, and Matthew C. Waxman. "Debating Autonomous Weapon Systems, their Ethics, and their Regulation under international law." (2017).

Williams, Andrew. "Defining Autonomy in Systems: Challenges and Solutions." *Issues for Defence Policymakers*(2015): 27.

Davison, Neil. "A legal perspective: Autonomous weapon systems under international humanitarian law." *Perspectives on lethal autonomous weapon systems* (2017): 5-18.

Scott, Ben, Stefan Heumann, and Philippe Lorenz. "Artificial Intelligence and Foreign Policy." *Stiftung Neue Verantwortung Policy Brief* (2018).

Lucas Jr, George R. "Automated Warfare." *Stan. L. & Pol'y Rev.* 25 (2014): 317.