National Defence Défense nationale

Canadian Forces College

Collège des Forces Canadiennes

WARBOTS AND AUTONOMOUS WEAPON SYSTEM-CAUSED DISRUPTIONS: A SYNTHESIS OF PRIORITIES FOR NATIONAL DEFENCE

Lieutenant-Colonel Sébastien Gorelov

JCSP 49

Exercise Solo Flight

Disclaimer

Opinions expressed remain those of the author and do not represent Department of National Defence or Canadian Forces policy. This paper may not be used without written permission.

PCEMI n° 49

Exercice Solo Flight

Avertissement

Les opinons exprimées n'engagent que leurs auteurs et ne reflètent aucunement des politiques du Ministère de la Défense nationale ou des Forces canadiennes. Ce papier ne peut être reproduit sans autorisation écrite.

Canada

CANADIAN FORCES COLLEGE - COLLÈGE DES FORCES CANADIENNES

JCSP 49 - PCEMI n° 49
2022 - 2023

Exercise Solo Flight – Exercice Solo Flight

WARBOTS AND AUTONOMOUS WEAPON SYSTEM-CAUSED DISRUPTIONS: A SYNTHESIS OF PRIORITIES FOR NATIONAL DEFENCE

Lieutenant-Colonel Sébastien Gorelov

WARBOTS AND AUTONOMOUS WEAPON SYSTEM-CAUSED DISRUPTIONS: A SYNTHESIS OF PRIORITIES FOR NATIONAL DEFENCE

> *Once Azerbaijan won air supremacy, their airborne robotic systems hunted targets inside the designated strike zones, day and night, at machine speeds*.
>
> — John Antal, *Seven seconds to die*

INTRODUCTION

The statement above by author John Antal[1] triggers fear of Autonomous Weapons Systems (AWS) hunting humans in an apocalyptic fashion. Indeed, when Azerbaijan and Armenia fought in 2020, Armenian troops did die from robotic systems used in war, but in this statement the question of human control and agency is concealed. Public and decision-maker understanding of what human control means is required to inform all discussions about warbots and AWS.[2] The following discussion considers military drones and software driven robots at all levels of automation. Author Kenneth Payne refers to these weapon systems as warbots or "intelligent warfighting machines."[3] The discussion will only focus on hardware and software acting in the physical plane.

This paper will argue how the Department of National Defence (DND) can best tackle the challenges and exploit the opportunities created by emerging AWS technologies in order to defend Canada and its interests. It will occasionally refer to the Second Nagorno-Karabakh war of 2020 between Azerbaijan and Armenia to showcase the current rapidly changing technological advancements worldwide in warbots and AWS. In the first part of the paper, we define the problem and examine how the important transformation is not caused as much by new warbot autonomy, but by global availability of enabling technology. In the second half of the paper we use these deductions to identify five areas in which government investments today will pay off greatly.

PART 1: DEFINITIONS, PROBLEM AND FAULT LINE

What are AWS?

Technology and warfare have always been tightly linked. The information age brings a new level of conceptual complexity. Understanding of the emerging advancements in warbots requires nuanced appreciation of the underlying technologies.

---

[1] John F. Antal, *7 Seconds to Die: A Military Analysis of the Second Nagorno-Karabakh War and the Future of Warfighting*, 2022), 122.

[2] The idea is reflected in Canada's Defence Policy Strong, Secure, Engaged: "The Canadian Armed Forces is committed to maintaining appropriate human involvement in the use of military capabilities that can exert lethal force." Canada, Department of National Defence, Strong, Secure, Engaged (SSE) (Ottawa, CA: 2017), 73, last accessed 7 May 2023, https://www.canada.ca/en/department-national-defence/corporate/policies-standards/canada-defence-policy.html.

[3] Kenneth Payne, *I, Warbot: The Dawn of Artificially Intelligent Conflict* (Oxford: Oxford University Press, 2021), 20.

The Royal Canadian Air Force (RCAF) is acquiring the same type of Remote Piloted Aircraft Systems (RPAS) aircrafts as the Royal Air Force. The RPAS project staff noted that in the United Kingdom, the public needed to be reassured that this drone is not autonomous, it is a Human In-The-Loop (HITL) system.[4] In fact, direct human control is a must for the civilian regulators who need to certify the coexistence of such aircraft with airliners in the busy skies above southern Canada. But even such a fully HITL system must be equipped with an automatic fail-safe. A back-up mode, in case of communication failure, accidental or adversarial, must allow the aircraft to fly a preprogrammed path to safe landing.[5]

Artificial Intelligence has existed since the beginning of digital computing. The word intelligence is both useful and misleading. It has qualified the leading edge of computing achievement throughout the years. What was artificial intelligence in research laboratories in the 1970 is just considered digital computing today. The word intelligence often only qualifies the latest computing breakthrough in capability, and is sometimes abused for effect. Artificial General Intelligence (AGI) "is a subset of AI whose proponents support the idea of the creation of a human-level intelligence". If regular AI applies to a defined problem and is thus narrow, AGI is much broader.[6] AGI is seen as an aspirational goal for future research.

It is important to understand the technical concept underlying Machine Learning (ML), a subset of AI, in order to not anthropomorphize the concept of learning. As opposed to the classic software approach of developing governing rules to be strictly followed, a designer using ML builds the software so it can refine its output using statistical methods. In other words, the software is designed to tune itself on a training set of data; human involvement at this step can vary between methods. The concept of "tuning" is key because the resulting process is fundamentally probabilistic.[7] At the two ends of the spectrum, AI now encompasses both deterministic systems at one end, where the outcome can be analyzed, bounded and understood, and stochastic systems at the other end, where the outcome can only be understood by random probability distributions. Speech recognition and computer vision are common applications that use ML. Although ML enabled impressive feats of engineering, the technology comes with important nonintuitive limitations that matter very much to AWS.

---

[4] "The committee's report observed a "sense of public disquiet" around the use of RPAS in military operations." Louisa Brooke-Holland, Overview of Military Drones used by the UK Armed Forces (London: House of Commons Library, 2015), 10.

[5] Some drones also have the ability to automatically temporally deviate from course to avoid another airplane if they sense an imminent collision. This system is called Sense and Avoid and may include a Due Regard Radar.

[6] D. Bruckner, H. Zeilinger and D. Dietrich, "Cognitive Automation-Survey of Novel Artificial General Intelligence Methods for the Automation of Human Technical Environments," *IEEE Transactions on Industrial Informatics* 8, no. 2 (2012), 208.

[7] "What is machine learning?" IBM, last accessed May 7, 2023, https://www.ibm.com/topics/machine-learning.

AWS are difficult to define. "Without a common lexicon, countries can have heated disagreements talking about completely different things."[8] At the state level, the international conversation on this topic started with the UN Convention on Certain Conventional Weapons (CCW) Meeting of High Contracting Parties with informal Meetings of Experts in 2014, 2015 and 2016. High Contracting Parties established a Group of Governmental Experts (GGE) on Lethal Autonomous Weapon Systems (LAWS).[9] As explained by Sauer, achieving clarity in the discussion about AWS is a struggle.[10] That is because of the potential arms control nature of the interstate conversation on one hand and the complexity and fluidity of the technology on the other.[11] For this paper we use Sauer's definition. AWS are "weapons capable of selecting and engaging targets without human intervention."[12] We further use Scharre's nomenclature to distinguish two important categories of AWS designs.[13] First, the concept of *supervised* autonomous weapons corresponds to Human On-The-Loop (HOTL) systems, which means humans can intervene in real-time to stop an engagement. An often-cited example in the literature is the Aegis naval defense systems. When saturation enemy attacks overwhelm the operator, the ship can defend itself at machine speed while allowing operator veto at any time. Second, the concept of *fully* autonomous weapons corresponds to Human Out-Of-The-Loop (HOOTL) systems, which once activated can function without human intervention. A low technology example of a fully autonomous weapon is a mine. A current high technology example is the HARPY Israeli drone. The Harpy is an Anti-Radiation homing kamikaze drone. It's more effectively described as a Loitering Munition (LM) version of the more classic Anti-Radiation missile technology that has existed since the US-Vietnam war. The drone identifies, self-selects and homes in on target Electro-Magnetic (EM) emissions. This is the kind of warbot that made the headlines during the war between Azerbaijan and Armenia in 2020.

Note that both supervised and full autonomy definitions use the verb *can*, meaning that there is potentially an overlap. This is an example of the fluidity of the technology. There could be an AWS that offers both supervised and full autonomy via software. For instance, humans could rely on a remote-control mission abort functionality but only up to certain conditions. For instance, *if* both the stakes are high enough (reward exceeds risk) *and* communications are lost due to enemy action, *then* humans would require the AWS to be fully autonomous. Usually, however, these two categories of supervised and full autonomy are kept distinct based on the designer's intent on how the

---

[8] Paul Scharre, *Army of None: Autonomous Weapons and the Future of War*, 1st ed. (London; New York: W.W. Norton & Company, 2018), 347.

[9] "Background on LAWS in the CCW – UNODA," United Nations Office for Disarmament Affairs, last accessed May 7, 2023, https://www.un.org/disarmament/the-convention-on-certain-conventional-weapons/background-on-laws-in-the-ccw/.

[10] Frank Sauer, "Autonomy in Weapons Systems and the Struggle for Regulation," Center for International Governance Innovation, November 28, 2022, 2. Last accessed May 7, 2023, https://www.cigionline.org/articles/autonomy-in-weapons-systems-and-the-struggle-for-regulation/.

[11] The UN definitions are included under CCW/GGE.1/2023/CRP.1 at https://meetings.unoda.org/meeting/67246/documents.

[12] Sauer, "Struggle for Regulation". Note that there are other definitions that can "shift the lexicon dramatically". In 2020 the UK doctrinal definition includes machines that "understand higher-level intent." in Scharre, "Autonomous Weapons and Stability," 96.

[13] Scharre, *Army of None*, 46.

system should normally be used. Inside the fully autonomous category, the important discriminating factors are what are the constraints on autonomy, or, decision-making freedom. The anti-personnel land mine is fixed in space but free to act indefinitely. In contrast, the Harpy has freedom to maneuver inside a human programmed window of time and space. However, it is likely that the Harpy is somewhat restricted in choosing the types of threats it can engage.[14] Finally, an expression that is subject to much debate is Meaningful Human Control (MHC). MHC generally gathers the ideas of ethical judgement, safety and controllability.[15] What emerges from the prolonged international discussions on autonomous weapons is the fundamental importance of MHC in understanding the future strategic impact of AWS.

Problem Definition

A review of more than 20 recent Canadian Forces College Joint Command and Staff Programme papers presents a noteworthy range of different points of view on how to best handle the evolving situation. Officers argue that the AWS revolution is upon us, others that it is not,[16] and that AWS should be embraced or banned.[17] The majority of papers argue that AWS should be adopted noting that technology outpaces regulation and that because of legal and ethical issues, Canada should approach the matter slowly and with caution. What can we make of all this? What should be the best approaches for Canada to adopt? The contribution of this paper is to synthesize and justify priorities on what should matter right now to Canadian government decision-makers.

In the book *Seven Seconds to Die*, John Antal depicts the overwhelming advantage gained by Azerbaijan over Armenia in 2020 by exploiting drone technology, both Remotely Piloted Uninhabited Combat Air Vehicles (UCAV), which are HITL, and Loitering Munitions with different degrees of autonomy, from HITL automation to supervised autonomy to HOOTL autonomy. The book is meant to be a wake-up call for Western allies to adapt quickly to this reality or lose their conventional military advantages to more agile nations. This idea is of strategic importance. However, misunderstandings about the exact nature of the underlying levels of automation and autonomy that were used by Azerbaijan lead to confusion. The media often broaches the topic from a sensationalizing narrative. This overdramatization, although understandable, risks clouding the judgement of the public, politicians and decision-makers. Cummings

---

[14] The Harpy decision-making freedom to engage targets is limited by its threat library and its identification capability. I speculate this is by design. The following Canadian Forces College paper contains a good list of current LAWS in service, including many discussed in this paper. See Appendix 2 in Daniel E. Hogan, "Sleepwalking into a Brave New World: The Implications of Lethal Autonomous Weapon Systems," (Directed Research Project, Canadian Forces College, 2021), 114.

[15] Paul Scharre, "Autonomous Weapons and Stability" (PhD, King's College London, 2020), 26.

[16] "Moving forward, the underlying need to maintain human control will limit the "decision making" ability of the machines used in war. So, although autonomous weapon systems will continue to evolve, they will not revolutionize the way that wars are fought." I agree with Owens in the sense that, for the reason stated, the revolution of AWS is not upon us. The current transformation in warfare is the proliferation of lesser autonomous systems. Gregory Owens, "Controlled Autonomy: The Limited Future Use of Autonomous Weapons," (Canadian Forces College, 2019), 12.

[17] Daniel Rice, "Lethal Autonomous Weapon Systems: A Clear and Present Danger," (Canadian Forces College, 2019), 15 and Hogan, "Sleepwalking into a Brave New World," 104.

writes that weapon systems of today "are more automated than autonomous. [. . .] This seemingly nuanced description is far from trivial and is critical for the debate about future lethal autonomous systems."[18]

The focus of much public discussions and awareness is centered too far on the full autonomy side. The reality right now, the true paradigm shift is not HOOTL but growing capabilities and great proliferation of lesser levels of autonomy, HOTL and below. Massive operationalization of full autonomy would be revolutionary to warfare but it is yet in the distant future. Common understanding needs to shift back towards the imminent threats and opportunities for Canada. The aim of this paper is to bring clarity to the topic and synthesize recommendations.

Fault Line

Warbots and AWS can be qualified by two independent variables: the level of direct human control and the decision-making freedom of the warbot. It is helpful to map this out: Figure 1 represents the two-dimensional space discussed here. Does one understand where new AWS systems exist in this space?

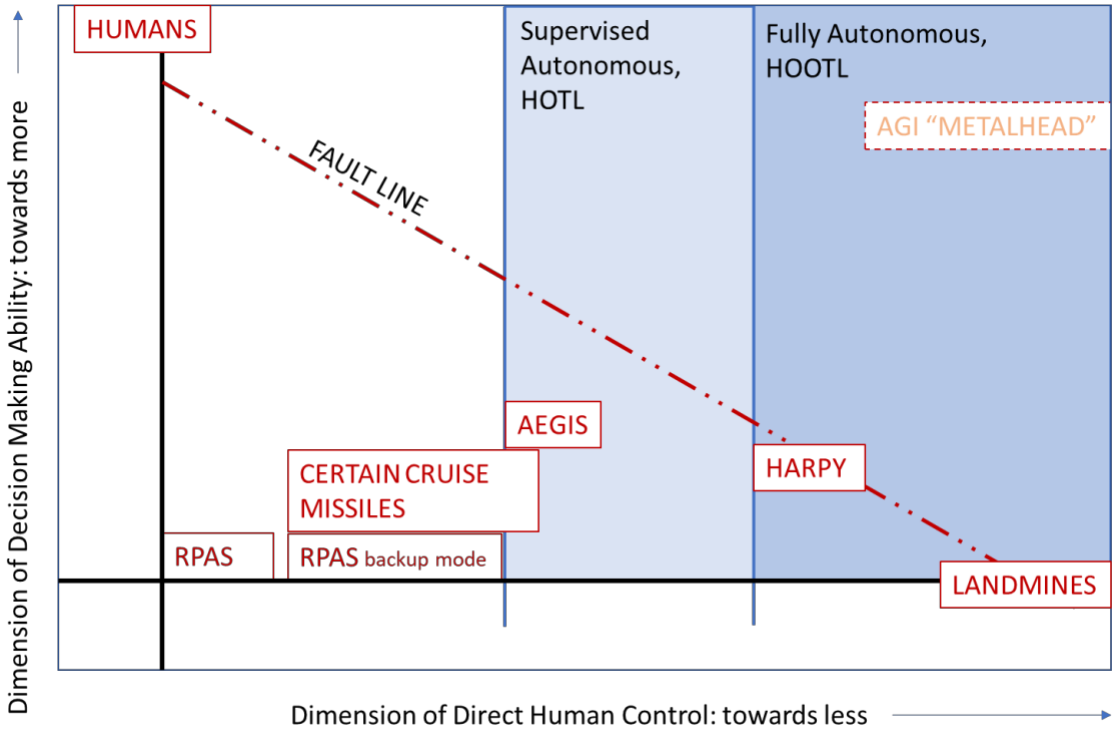

Figure 1 – A warbots' level of direct human control versus its decision-making ability

[18] M. Cummings, "The Human Role in Autonomous Weapon Design and Deployment," in *Lethal Autonomous Weapons: Re-Examining the Law and Ethics of Robotic Warfare*, ed. Jai Galliott, Duncan MacIntosh, and Jens David Ohlin (Oxford: Oxford University Press, 2021), 277.

We will see that the immediate challenge does not lie in the right side of the chart, but on its bottom half. The fault line represents a clear gap between where AWS are proliferating now and where advances will lead us in the future. It is currently extremely difficult for reliable new designs to cross that line towards more decision-making freedom as we examine below. Confusion between the two sides of the line complicates the debate on what should be done to adapt to the new reality. We now look at both of the independent variables.

On the Horizontal Axis, Direct Human Control

Human Supervisory Control (HSC) is an important engineering field of research that has numerous civilian applications, from nuclear power plant, spacecraft to aviation applications. A leader in the HSC field, Cummings recognizes that "it is often hard to draw the line between automated and autonomous." [19] She describes how designers balance "authority between the human and the computer." In civilian applications, "many control engineers see the human as a mere disturbance." This is because civilian applications have a narrow and routine scope and different operators bring a wide-ranging set of different levels of performance. "Rules and criteria […] reduce the ambiguity in the design space."[20] Self-driving cars are an example of successful civilian application of autonomous systems based on ML technologies. In self-driving, the autonomy feature acts as a relief, HOTL type system. It works well enough based on the codified, ruled-based, visual environment of the road traffic system. Although self-driving autonomy does improve over time, it fails quickly in atypical conditions. "The inability of such algorithms to cope with uncertainty in autonomous systems is known as brittleness, which is a fundamental problem for computer vision based on deep learning."[21] Scharre confirms this in his work. Atypical situations and actively adversarial actors are wartime conditions that "undermine reliability"[22] for fully autonomous weapons.

On the Vertical Axis, Decision-Making Freedom[23]

Cummings' research highlights the limitations of computers in situations when human expertise is still the best at handling uncertainty in complex environments[24]. Providing a computer freedom to make decisions is easy, but designing a system which can reliably make good decisions, let alone surpass expert human operators, becomes

---

[19] Cummings, "The Human Role," 274.
[20] Cummings, "The Human Role," 275.
[21] Cummings, "The Human Role," 281.
[22] Scharre, "Autonomous Weapons and Stability," 4.
[23] Payne differentiates human biological agency from machine agency. Here I deliberately avoid the anthropomorphizing term *agency*, even though the machines "can choose" to a certain extent. In Payne, *I, Warbot*, 24.
[24] "Because of the aforementioned brittleness problems in the programming of computer algorithms and the inability to replicate the intangible concept of intuition, knowledge-based reasoning, and true expertise, for now, are outside the realm of computers. However, there is currently significant research underway to change this, particularly in the machine learning (sometimes called artificial intelligence) community, but progress is slow." In Cummings, "The Human Role in Autonomous Weapon Design and Deployment," 283.

exponentially harder as the uncertainty grows. In the science-fiction television streaming series *Black Mirror*, an episode called *Metalhead* features a lethal robot dog. This warbot offers an example of a fully autonomous LAWS with extraordinary level of decision-making freedom, which reveals its benefits as a relentless human killing machine. This is a useful fictional example of the nightmarish vision of LAWS which can cloud the debate: How fictional is it after all? Right now, the most likely path towards an autonomous warbot displaying such levels of decision-making freedom would be that of Artificial Neural Networks, a more sophisticated subset of ML. This path will have to overcome tremendous obstacles to get to the level of decision-making freedom featured in the *Black Mirror* episode. In the book Army of None, Scharre lists the challenges with AWS driven by ML technologies, all due to their fundamentally stochastic nature. The main points are synthesized here:

Comprehensibility: Due to the "black box" nature of the design it is difficult to understand what and why the system reacts this way to inputs. Is it acting beyond our intent?[25]

Trust: Explosives can only be operationalized if fuse technology offers trust that the destructive effect is unleashed strictly on command, and if imperfect, it can at least be bounded, or fails safe. Related to this idea is risk assessment: How do you test and evaluate that a system can function as intended in the real adversarial world of combat when it is trained, at best, in simulated combat?[26]

Biases and "perverse instantiation":[27] How do you faithfully meet human intent, which may very well evolve over time depending on the circumstances? How does the system handle differences between designer and operator intent?

To summarize, designing for appropriate control is very challenging and is a requirement for powerful actors. Fielding of weapons without solutions to these issues would be rather easy to counter, providing a damper to rapid AWS operationalization. Because of brittleness, the immediate future looks more like a version of the *Metalhead* warbot that is remote-controlled by humans rather than the fully autonomous warbot depicted in the film. That is still worrisome. As highlighted in Cummings's work, the combination of both machine and human strengths is contributing to the changing reality.

Seven seconds to die describes the conflict in Nagorno-Karabakh as being one of unmanned systems:

> *Most of the casualties inflicted upon the Armenians were from TB2 UCAVs that launched smart micro-munitions, the HAROP "kamikaze" LMs, and*

---

[25] Scharre, *Army of None*, 180.
[26] Scharre, *Army of None*, 149.
[27] Nick Bostrom, *Superintelligence: Paths, Dangers, Strategies* (Oxford: Oxford University Press, 2014), Chapter 8 as quoted in Scharre, "Autonomous Weapons and Stability," 236.

> *other UAVs that designated targets for UCAVs, LMs, and long-range artillery*.[28]

Indeed, the drones that caused the most decisive damage were not HOOTL systems[29]. The transformation in warfare is happening on the automatic and remotely controlled end of the autonomy spectrum, not on the fully autonomous side. The center of gravity of the strategic discussion should therefore shift towards semi-autonomous and remote-controlled systems which is currently the biggest novel threat. Full autonomy weapons are currently too brittle and may be threatening later in the future. Key investments today should reflect this reality and we identify five priority areas divided in two groups. The first group of priorities addresses the lesser autonomous systems. Investments in defensive and offensive systems are important; for the former they are indeed underway. Human Machine Interaction research belongs in the middle, as we have seen the preeminence of human control. The second group addresses concern with the more autonomous systems. Here the key efforts should be both investments in counter AWS technologies and from a strategic, longer term point of view, continuous participation in the regulation debate at the global level. We synthesize those priorities below.

PART 2: PRIORITIES FOR DND

Effectiveness for a Fraction of the Cost: Defend against Warbot Proliferation:

> "Russian forces […]in Syria came under attack from mini-swarms of 10 and 3 'small' strike UAV respectively. [. . .] The munitions used were described as having been 'improvised explosive devices' that were releasable. [. . .] These attacks may be considered as potential future rogue weapons, and [. . .] set a rather chilling precedent for the types of weapons that terrorist organizations can increasingly produce in-house."[30]

This attack happened in January 2018 and proves that the proliferation of robotic threats is not in the future, indeed it has already happened. Proliferation of such adversarial capability is the biggest AWS threat to Canada today. This is true even if the level of autonomy and decision-making freedom of the robotic systems is very low, for instance with preprogrammed kamikaze drones, unable to adapt to changes in environmental conditions in real-time. Both the availability and affordability of the technology explain the proliferation. Commercially available hardware and software

---

[28] Antal, *7 Seconds to Die*, 126.

[29] The HAROP can feature different levels of autonomy above and below HOTL but is advertised as HOTL. Furthermore, "today, we cannot replicate the human brain. We cannot develop a machine that has the decision-making abilities, the ethical values, and the morals of a human. Someday, this may be possible, but most likely not for a while." In Antal, *7 Seconds to Die*, 123.

[30] Martin Streetly and Beatrice Bernardi, *Jane's all the World's Aircraft: Unmanned: 2018-2019* IHS Jane's, 2018), 6 and "Syria war: Russia thwarts drone attack on Hmeimim airbase," *BBC*, January 7, 2018, https://www.bbc.com/news/world-europe-42595184.

result in non-state actors and individuals having access to low-cost remotely operated drones and software that can greatly enhance the capabilities of these drones.[31]

A review of Janes 2018-2019 Unmanned aerial vehicles confirms proliferation in state arsenals as well. The permanent five United Nation Security Council members are all investing in pushing their robotic drone capabilities. Countries with humbler means are following suit and sometimes leading the way: Israel, Turkey and Iran are dominating segments of the market. The unmanned arms industry is booming and actors can acquire capabilities proportional to their budgets or develop them themselves. To provoke South Korea, North Korea flew its own drones over Seoul in December 2022.[32] According to the article, it did not cause any physical damage. In contrast, multiple attacks against Saudi oil infrastructure have happened since an attack in September 2019, temporarily stopping oil processing and causing a sharp jump in crude futures.[33] What is happening worldwide is therefore a gradual closure in the conventional warfare gap.[34] This conventional gap between well-funded militaries and underfunded forces is closing at a speed that is accelerating. This was apparent in Nagorno-Karabakh, where the Russian-supported forces in Armenia lost air superiority quickly, and this is somewhat apparent in Ukraine, where the might of the Russian military was unable to break Ukrainian defenses in large parts of the front.

The Canadian Army understands this threat and is adapting by rethinking its force protection. It is leading a positive push in counter Uncrewed Aircraft System (UAS) and air defence procurement, with timeline and funding issues still under discussion.[35] In parallel the RCAF should not delay investment in Ground Based Air Defence specifically to counter the drone threat. To counter such threats effectively, Canada should be ready to employ defensive AWS to compete in "machine-time". Indeed, "without autonomous modes of operation, human operators could be overwhelmed by short warning saturation attacks."[36]

The proliferation of adversary automated weapons in the form of robotic drones is more of a threat in itself than the autonomous nature of the drones. It is the availability and affordability of this technology, more than its autonomous capability, that decreases

---

[31] Scharre, "Autonomous Weapons and Stability" , 114.
[32] Jean Mackenzie, Robert Plummer, "North Korean drone reaches north of Seoul," *BBC*, December 26, 2022, https://www.bbc.com/news/world-asia-64094143.
[33] "Yemen's Houthis claim drone attack on refinery in Saudi capital," *Reuters*, March 11, 2022, https://www.reuters.com/world/middle-east/attack-refinery-riyadh-did-not-affect-petroleum-supplies-spa-2022-03-10/ and "Oil prices spike after Saudi drone attack causes biggest disruption ever – as it happened," *The Guardian*, September 16, 2019, https://www.theguardian.com/business/live/2019/sep/16/oil-price-saudi-arabia-iran-drone-markets-ftse-pound-brexit-business-live.
[34] "AI and robotics will smash the status quo that exists in the world today and will reduce the gap between advanced military powers and the rest of the world." Abishur Prakash quoted in Antal, *7 Seconds to Die: A Military Analysis of the Second Nagorno-Karabakh War and the Future of Warfighting*, 132.
[35] Canada, Department of National Defence, Defence Capability Board Record of Discussion. (Ottawa, CA: 2022).
[36] Scharre, "Autonomous Weapons and Stability," 20.

the gap in conventional advantage that DND use to enjoy. What opportunities might this same dynamic present to Canada?

Effectiveness for a Fraction of the Cost: Opportunities

The same availability and affordability logic applies to the DND as well. Canada can effectively employ drone technology to reduce the cost of airpower. A recent news article claims that in October 2020, a Harop (Harpy-2) Israeli built kamikaze drone from Azerbaijan successfully destroyed a Russian-build S300 Surface-to-Air missile system from Armenia.[37] Janes 2018-2019 suggests that a Harpy anti-radar kamikaze UAV is USD450k$.[38] The Harpy 2 is newer that the Harpy, features longer range endurance and is recoverable. Let's guess that a Harop is at most 10 times more expensive, at USD5M$ which is probably a reasonable overestimated limit. If the article is accurate, the drone was able to prosecute an S300 site estimated at USD50M$.[39] For this argument, this is a ten-to-one target to munition cost ratio at worst and multiple times that at best. Consider what conventional strike fighter resources are required to obtain the same effect, plus the operational risk involved. The Suppression of Enemy Air Defense (SEAD) mission risk-reward calculus is fundamentally different. Thus, the paradigm shift is the availability of effective airpower at a fraction of the cost. It is possible that with the F-35 procurement, the RCAF will peak in term of cost per platform, and that in the future, such capability will be too impractical to sustain. Realization of this change is spreading around the world. Russia is also seizing the opportunity when using Iranian made drones in Ukraine, taking advantage of their relative low cost.[40]

This transformation is an opportunity for Canada. The RCAF is preparing to operate the MQ9B armed surveillance RPAS. The Turkish-build TB2 can do equivalent missions sets, but is not as high performance as the MQ9B. Without speculating on the qualitative difference, large cost differences open up opportunities to rethink airpower and pan-domain force employment. If a country can afford ten time as many TB2s as MQ9Bs, should it invest in numbers over quality?[41] In this particular example, Canada is better off with the MQ9B due to arctic surveillance requirements. The question is still worth asking in a larger sense.[42] RPAS risk tolerance in a theater of war is different than RPAS risk tolerance in domestic airspace shared with civilian airliners. On the battlefield, survivability is an issue, for MQ9B and TB2 equally. UAS "must be cheap and available

---

[37] "Azerbaijan used Harpy-2 drone to destroy another S-300 SAM site in Armenia," Global Defense Corp, October 26, 2020, https://www.globaldefensecorp.com/2020/10/26/azerbaijan-used-harpy-2-drone-to-destroy-another-s-300-sam-site-in-armenia/.

[38] Streetly, *Jane's all the World's Aircraft: Unmanned: 2018-2019*, 110.

[39] Estimated as follows: An S300 battery was USD150M$ in 2010 and has three launcher trucks plus one radar. See Dmitry Solovyov, "China buys air defense systems from Russia," Reuters, April 2, 2010, https://www.reuters.com/article/us-russia-china-arms-idUSTRE6310WG20100402.

[40] Chris Gordon, "Cheap UAVs Exact High Costs," *Air and Space Forces Magazine*, Jan 20, 2023, https://www.airandspaceforces.com/article/cheap-uavs-exact-high-costs/.

[41] John Antal advances a figure between USD1 and 2M$ per TB2. That would mean a 20-to-1 ratio between the two drones. Antal, *7 Seconds to Die*, 128.

[42] "Investments and trials in cheaper and similar systems could satisfy our nation's procurement system's risk-averse nature" in Michael J. Ulloa, "The Effects Of Unmanned And Autonomous Weapons," (Canadian Forces College, 2021), 7.

in quantity", as noted in the initial lessons learned from the war in Ukraine. Furthermore, "UAS and CUAS must be available across all branches and echelons."[43]. If the Canadian Armed Forces (CAF) wants to survive on the battlefield, it should pursue both high and low-cost solutions.

Finally, this reality is a fundamental change for security and stability on the global stage. It would be a mistake to dismiss acquiring lesser capable systems because of domestic regulator issues, like with Transport Canada imposing requirements and constraints that do not provide an advantage in the theater of operation. As for the more fully autonomous systems, Bradley Perry draws attention to the benefits of autonomous convoys in a 2021 Canadian Forces College paper.[44] There are indeed tremendous opportunities to employ autonomous systems that are not weapons inside the CAF. In order to capably operationalize DND drones and autonomous systems it should consider procuring, the government of Canada should invest in Human-Machine Interaction research.

Human-Machine Interaction (HMI)

Human Factors in Aviation (HFA) is a discipline that emerged from World War II in parallel with ergonomics research which, in a broader sense, tackled industry wide technical challenges for worker productivity. HFA is relevant here because aviation is where technology and pilot vehicle interface has had a net positive impact on safety in the civilian sector. HFA evolved to include "usability, training, design, maintenance, safety, procedures, communications, workload and automation." It also evolved to include Crew Resource Management (CRM) to emphasize teamwork. The same idea emerges for Human Machine Teaming (HMT) when the computer's capability increases to the point that it is promoted to be a part of the team as an "agent." Literature on HMT differentiates it from HMI based on the level of agency of the machine.

For HMI in military aviation cockpits, Pilot Vehicle Interface (PVI) design is especially important for platforms that handle multiple overlapping systems, like when single seat fighter pilots manage sensors, weapons, defence systems, navigation and flight controls when the attention of the operator is the most critical resource. For fourth generation fighters, Hands on Throttle-and-Stick (HOTAS) "switchology" is an example of PVI that can make a difference in weapon system performance. In fact, differences in PVI designs are exploitable when designing combat tactics against specific known adversaries.

The problem of how to best arrange Humans and Machines working together is therefore not new. Two decades ago, when speaking of humans "sharing control of systems with automation" Leveson wrote:

> These changes are leading to new types of human errors and a new distribution of human errors (for example, increasing errors of omission

[43] Mykhaylo Zabrodskyi et al., *Preliminary Lessons in Conventional Warfighting from Russia's Invasion of Ukraine: February–July 2022* Royal United Services Institute,[2022])., 57-58.
[44] Bradley Perry, "Autonomous Convoys," Canadian Forces College, 2021.

> versus commission). All human behavior is influenced by the context in which it occurs, and operators in high-tech systems are often at the mercy of the design of the automation they use.[45]

On 29 April 2020, an RCAF Cyclone crashed in the Ionian Sea. The *Stalker-22* accident is an example of how industry and operators are still learning how to master automation, even for deterministic implementations.[46]

In the sense that mature technologies still lead to accidents, AI and AWS technologies are not bringing a revolution. It is already difficult to master automation with Humans In-The-Loop in deterministic systems. The future of airworthiness and operational suitability is, more than ever before, about human control. As automation progresses towards autonomy, several problems are exacerbated and we should keep using the methods and knowledge gained to progress in that continuum. "As automated systems become more sophisticated, human factors considerations become more vital to address."[47]

We saw how the possible evolution from deterministic to stochastic systems creates new problems related to the probabilistic nature of software. "Perverse Instantiation" is when the imperfect communication of requirement by humans leads the machine to a destructive outcome. This led Kenneth Payne in *I, Warbot* to formulate this proposed law of Lethal AWS: "A warbot should understand my intentions, and work creatively to achieve them." This would also help with the problem of "projection bias" where humans "incorrectly project their current beliefs and desires onto others and even their future selves." This concept of the machine "checking-in" to understand human intent is emerging more generally as a measure of safety for AGI development.[48] HMI research can find out how to best implement this idea.

It appears that for the foreseeable future, humans and machines continue to bring complementary capabilities to the fight. "The best weapon systems would be those that optimally use both humans and automation."[49] The government of Canada should invest in Human-Machine Interaction research. The optimization of the interface can yield a decisive advantage in combat. We examined priorities that concern the spectrum of AWS

---

[45] Mode confusion is one such type of new human errors. Nancy Leveson, "A New Accident Model for Engineering Safer Systems," *Safety Science* 42, no. 4 (2004), 239.

[46] An aircraft fly-by-wire control law design assumption by Sikorsky was hidden to the operators. Indeed, proper procedures masked the fact that in case of omission of a critical step, in certain conditions full pitch-up authority in the controls would be insufficient to recover quickly and safely from a dive at low altitude. It is just a matter of time for pilots to make procedural omissions. Designers and testers missed the critical significance of that step-in certain conditions. In this case the accident cost lives. Canada, Department of National Defence, CH148822 Flight Safety Investigation Report (Ottawa, CA: 2021), 27.

[47] Gemma J. M. Read et al., "What is Going on? Contributory Factors to Automation-Related Aviation Incidents and Accidents," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 64, no. 1 (2020), 1700.

[48] Max Tegmark interview, discussing Stuart Russel's Inverse Reinforcement Learning, in "The Case for Halting AI Development," April 13, 2023, *Lex Fridman Podcast 371*, produced by Lex Fridman, MP3 audio, 2:18:00, https://lexfridman.com/max-tegmark-3/.

[49] Scharre, "Autonomous Weapons and Stability," 242.

that involves less autonomy. Next, we look at efforts that the government of Canada should prioritize to tackle the challenge of future fully autonomous weapon systems.

Counter-AWS

What would happen should one group of countries responsibly regulate their operationalization of fully autonomous LAWS by following proposed rules, but not their adversary? As it turns out, it could be a routine matter of understanding and countering the threat. Just like the CAF possesses a credible counter Chemical, Biological, Radiological and Nuclear capability, it should acquire a similar counter AWS programme.

The Future of Life institute has created a short film in 2017 called Slaugtherbots.[50] In it, micro drone swarms, each carrying a lethal charge to attack skulls, are shown employed by terrorists to target student activists. The students are identified by their pictures and social media presence. The technology exists today to have a loitering munition be programmed to detect human beings inside a defined area and autonomously kill them, like a loitering anti-personnel air mine. What does not exist is the ability to distinguish legal targets, in the sense of the Law of Armed Conflict. Distinguishing legal targets from non-combatants with accuracy would fulfill the principle of discrimination. The film shows the terrifying weaponization of autonomous systems which use advanced neural nets to make decisions about humans as targets. NGO campaigns like this one have an important role to play, but the sensationalism of the narrative should be accompanied by constructive descriptions on advances in AI safety and counter autonomy research.

Target selection for example can be fooled if based on computer vision processing, as explained by Scharre.[51] The same technology that enables the statistical predicting ability of ML can be used to manipulate the decision-making process. Slight alteration in input, undetectable by humans, can produce dramatic and repeatable changes in output. Adversarial ML is a new field of research that has, among other things, produced image cloaking software, like the one called *Fawkes*:

> *At a high level, Fawkes "poisons" models that try to learn what you look like, by putting hidden changes into your photos, and using them as Trojan horses to deliver that poison to any facial recognition models of you.*[52]

---

[50] "Slaughterbots," directed by Stewart Sugg (Space Digital, 2017), 8 min. https://www.youtube.com/watch?v=9CO6M2HsoIA.

[51] Scharre, *Army of None*, 182.

[52] Shawn Shan and Emily Wenger, "Fawkes," University of Chicago Computer Science, last accessed May 8, 2023, https://sandlab.cs.uchicago.edu/fawkes/.

The changes are not detectable by the human eye. Scharre believes such vulnerability "casts doubt on the wisdom of using the current class of visual object recognition AIs for military applications."[53]

When trying to counter AWS, if you cannot exploit or defeat the decision-making algorithm, you can next defeat the sensors the system relies on. Classical electronic warfare tools are designed to do exactly that. The real question is the practicality and affordability of Counter AWS solutions. Is the countermeasure cheaper than the threat? Right now, the answer is yes, but investments in research in this field will provide situation awareness should this change substantially.

The Economist article about Slaughterbots predicts accurately that it is just a matter of time before "someone, somewhere"[54] produces a capable lethal AWS to target humans. Whatever the government of Canada chooses to do, unscrupulous adversaries will continue to exploit the advantages provided by fully autonomous weapons. Canada should thus invest in counter AWS Research and Development. This could be first led by Defence Research and Development Canada and then scaled up to a counter AWS programme later, depending of the future threat of such weapons. Counter AWS experts could contribute to ally efforts in that domain at the tactical level and provide recommendation to Canadian decision-makers at the strategic level. This ties in with the final topic that the government should address.

Contribute to Strategic Stability

At the strategic level, it is more difficult to understand what would happen should a major player ignore internationally developed norms. Scharre's doctorate thesis analyses the question of strategic stability and proposes courses of action. He points out that predictability and controllability are an issue and quotes Payne in that "unintended escalation" is a risk when AI is entrusted to make strategic decisions.[55] If the big powers cooperate and agree to implement and verify norms, the risks of unintended escalation would be mitigated. This would matter more in times of crisis, outside of war. Scharre's options include restricting antipersonnel lethality and enforcing rules like "returning fire must be limited, discriminating, and proportionate,"[56] with the goal of avoiding "unintended escalation" between warbots and between humans.

In 2019, Prime Minister Trudeau's mandate letter to the then Minister of Foreign Affairs, Mr. Champagne, required him to advance "international efforts to ban the development and use of fully autonomous weapons systems."[57] The statement disappeared from the 2021 Mandate letter. This is perhaps because of the ill-defined nature of the problem. It is a good initiative to walk back from a self-imposed ban. The

---

[53] Scharre, *Army of None*, 182.
[54] The Economist, "Military robots are getting smaller and more capable," Dec 14, 2017.
[55] Payne, *I, Warbot*, 29.
[56] Scharre, *Army of None*, 357.
[57] Canada, Prime Minister's Office, *Archived – Minister of Foreign Affairs Mandate Letter* (Ottawa, CA: 2019), last accessed 8 May 2023, https://pm.gc.ca/en/mandate-letters/2019/12/13/archived-minister-foreign-affairs-mandate-letter.

government of Canada should strengthen strategic and global level engagement on the international AWS regulation debate, vice pursuing an a-priori ban. The goal should remain on containing strategic escalation, and if possible, restricting anti-personnel applications of AWS.

The urgent strategic issue is the proliferation of cheap, precise and effective instruments of airpower, and as a consequence the reduction of the asymmetry gap between powerful militaries and others.

Conclusion

In his book, Scharre highlights the advantage of increasing autonomy in modern day warbots: speed on one hand and effectiveness under conditions of communications black out on the other. Because of brittleness however, machine autonomy and decision-making freedom can currently only go so far in the atypical environment that exist during warfare. Public debate needs to bring its attention on the real story: today's paradigm shift is not about advances in machine decision-making abilities but in the remarkable proliferation of low autonomy warbots. How should the government of Canada respond to the challenges of this disruption? By pursing the five avenues proposed above: First, the CAF should keep working to protect itself against drones by the adversary; their ease of access and capability is a serious threat. Second, the CAF should demonstrate agility in acquiring the equivalent low-cost capabilities and rethink its pan-domain application of airpower. Third, it should invest in Human Machine Interaction research and development. This is the technical path to maintaining appropriate human control and developing technology responsibly. Fourth, it should establish counter autonomy expertise, to be ready to exploit vulnerabilities from future adversary systems. Finally, it should participate and advocate for the development of responsible international norms in warbots and AWS.

BIBLIOGRAPHY

Antal, John F. *7 Seconds to Die: A Military Analysis of the Second Nagorno-Karabakh War and the Future of Warfighting* 2022.

BBC. "Syria war: Russia thwarts drone attack on Hmeimim airbase." January 7, 2018. Last accessed May 8, 2023. https://www.bbc.com/news/world-europe-42595184.

Bostrom, Nick. *Superintelligence: Paths, Dangers, Strategies*. Oxford: Oxford University Press, 2014.

Bruckner, D., H. Zeilinger, and D. Dietrich. "Cognitive Automation-Survey of Novel Artificial General Intelligence Methods for the Automation of Human Technical Environments." *IEEE Transactions on Industrial Informatics* 8, no. 2 (2012): 206-215.

Canada. Department of National Defence. *CH148822 Flight Safety Investigation Report*. Ottawa, CA: 2017.

Canada. Department of National Defence. *Defence Capability Board Record of Discussion*. Ottawa, CA: 2022.

Canada. Department of National Defence. *Strong, Secure, Engaged (SSE)*. Ottawa, CA: 2017. Last accessed 7 May 2023. https://www.canada.ca/en/department-national-defence/corporate/policies-standards/canada-defence-policy.html

Canada. Prime Minister's Office. *Archived – Minister of Foreign Affairs Mandate Letter*. Ottawa, CA: 2019. Last accessed 8 May 2023. https://pm.gc.ca/en/mandate-letters/2019/12/13/archived-minister-foreign-affairs-mandate-letter.

Cummings, Missy, "The Human Role in Autonomous Weapon Design and Deployment." In *Lethal Autonomous Weapons: Re-Examining the Law and Ethics of Robotic Warfare*, edited by Jai Galliott, Duncan MacIntosh, and Jens David Ohlin, 273-287. Oxford: Oxford University Press, 2021.

Global Defense Corp. "Azerbaijan used Harpy-2 drone to destroy another S-300 SAM site in Armenia." October 26, 2020. Last accessed May 8, 2023. https://www.globaldefensecorp.com/2020/10/26/azerbaijan-used-harpy-2-drone-to-destroy-another-s-300-sam-site-in-armenia/.

Gordon, Chris. "Cheap UAVs Exact High Costs." *Air and Space Forces Magazine*. Jan 20, 2023. Last accessed May 8, 2023. https://www.airandspaceforces.com/article/cheap-uavs-exact-high-costs/.

Hogan, Daniel E. "Sleepwalking Into a Brave New World: The Implications of Lethal Autonomous Weapon Systems." Directed Research Project, Canadian Forces College, 2021.

IBM. “What is Machine Learning?” Last accessed May 7, 2023. https://www.ibm.com/topics/machine-learning.

Leveson, Nancy. "A New Accident Model for Engineering Safer Systems." *Safety Science* 42, no. 4 (2004): 237-270.

Mackenzie, Jean and Plummer, Robert. “North Korean drone reaches north of Seoul.” *BBC*. December 26, 2022. Last accessed May 8, 2023. https://www.bbc.com/news/world-asia-64094143.

Owens, Gregory. “Controlled Autonomy: The Limited Future Use of Autonomous Weapons.” Canadian Forces College, 2019.

Payne, Kenneth. *I, Warbot: The Dawn of Artificially Intelligent Conflict*. Oxford: Oxford University Press, 2021.

Perry, Bradley. “Autonomous Convoys.” Canadian Forces College, 2021.

Read, Gemma J. M., Alison O’Brien, Neville A. Stanton, and Paul M. Salmon. "What is Going on? Contributory Factors to Automation-Related Aviation Incidents and Accidents." *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 64, no. 1 (2020): 1697-1701.

Reuters. “Yemen’s Houthis claim drone attack on refinery in Saudi capital.” March 11, 2022. Last accessed May 8, 2023. https://www.reuters.com/world/middle-east/attack-refinery-riyadh-did-not-affect-petroleum-supplies-spa-2022-03-10/.

Rice, Daniel. “Lethal Autonomous Weapon Systems: A Clear and Present Danger.” Canadian Forces College, 2019.

Sauer, Frank. Center for International Governance Innovation. “Autonomy in Weapons Systems and the Struggle for Regulation.” November 28, 2022. Last accessed May 7, 2023. https://www.cigionline.org/articles/autonomy-in-weapons-systems-and-the-struggle-for-regulation/.

Scharre, Paul. *Army of None: Autonomous Weapons and the Future of War*. 1st ed. London; New York: W.W. Norton & Company, 2018.

Scharre, Paul. "Autonomous Weapons and Stability." PhD diss., King's College London, 2020.

Shan, Shawn and Wenger, Emily. “Fawkes.” University of Chicago Computer Science. Last accessed May 8, 2023. https://sandlab.cs.uchicago.edu/fawkes/.

Solovyov, Dmitry. “China buys air defense systems from Russia.” *Reuters*. April 2, 2010. Last accessed May 8, 2023. https://www.reuters.com/article/us-russia-china-arms-idUSTRE6310WG20100402.

Streetly, Martin and Beatrice Bernardi. *Jane's all the World's Aircraft: Unmanned: 2018-2019* IHS Jane's, 2018.

Sugg, Stewart, dir. “Slaughterbots.” Space Digital, 2017, 8 min. https://www.youtube.com/watch?v=9CO6M2HsoIA.

Tegmark, Max. “The Case for Halting AI Development.” *Lex Fridman Podcast 371*. April 13, 2023. Lex Fridman, MP3 audio. Last accessed May 7, 2023. https://lexfridman.com/max-tegmark-3/.

The Economist. “Military robots are getting smaller and more capable,” Dec 14, 2017.

The Guardian. “Oil prices spike after Saudi drone attack causes biggest disruption ever – as it happened.” September 16, 2019. Last accessed May 8, 2023. https://www.theguardian.com/business/live/2019/sep/16/oil-price-saudi-arabia-iran-drone-markets-ftse-pound-brexit-business-live.

Ulloa, Michael J. “The Effects Of Unmanned And Autonomous Weapons.” Canadian Forces College, 2021.

United Nations Office for Disarmament Affairs. “Background on LAWS in the CCW – UNODA.” Last accessed May 7, 2023. https://www.un.org/disarmament/the-convention-on-certain-conventional-weapons/background-on-laws-in-the-ccw/.

Zabrodskyi, Mykhaylo, Jack Watling, Oleksandr V. Danylyuk, and Nick Reynolds. *Preliminary Lessons in Conventional Warfighting from Russia’s Invasion of Ukraine: February–July 2022*: Royal United Services Institute, 2022.