ARE LETHAL AUTONOMOUS WEAPONS SAFE? EVALUATING THE CONCERNS AND BENEFITS OF KILLER ROBOTS

FERRARA, ZAIDA
Psychology and Neuroscience, College of Arts and Humanities

ABSTRACT

The advancement of technology in the military domain has the potential to position robotic warfare as a replacement for nuclear weapons. An emerging issue surrounding the use of lethal autonomous weapons (LAWs) is weighing the balance between reduced psychological trauma for military personnel and the potential for increased psychological terror of civilians in war zones. Supporters of robotic warfare argue that these systems offer significant advantages, including the ability to avoid the ethical and emotional challenges faced by human soldiers. Specifically, they assert that LAWs can mitigate the psychological toll of warfare, such as post-traumatic stress disorder and other combat-related trauma. However, critics of robotic warfare highlight the dangers of reduced accountability and the loss of ethical control in war. While using LAWs might lower the risk for soldiers, it could make governments more likely to start conflicts if they view these weapons as a safer option. History shows that military technology often focuses on improving efficiency, making it likely that autonomous systems will eventually replace traditional methods. Given these considerations, it is essential to carefully evaluate the advantages and disadvantages of LAWs, as this review seeks to do. To understand the ethical concerns surrounding this topic, it is necessary to engage in an interdisciplinary discourse, as policymakers, military officials, and psychologists must carefully consider whether the pursuit of LAWs serves the long-term interests of society or severely fuels mistrust between nations. Furthermore, the substitution of human judgment with machine decision-making in warfare also raises human rights concerns that require thorough examination. These issues must be addressed before any consideration is given to the large-scale deployment of LAWs, as they pose critical ethical and legal challenges.

INTRODUCTION

As technology progressively penetrates our daily lives, advanced military weapons are increasingly automated, evolving from smart munitions which autonomously locate targets to drones and robots capable of acting with minimal human intervention (Woods, 2020). In recent years, Lethal Autonomous Weapons (LAWs) have garnered attention within social robotics and Artificial Intelligence (AI) fields (Michalec and O'Donovan, 2021). The military is one of the largest funders of robotics research, driving innovations that make such technology ever more prevalent in warfare (Michalec and O'Donovan, 2021). Proponents, including military personnel, argue that these technologies could reduce civilian casualties and military fatalities, making battlefields less deadly overall (Woods, 2020). While current automated weapons generally operate under some level of human oversight, the human input required is minimal. The ultimate goal of many robotics developers is to create autonomous weapons that function independently, making complex decisions without human direction (Kormushev, 2013). These machines would become fully task-driven, acting purely on programmed objectives rather than human judgment or ethical considerations.

The rationale of this review is to present a summary of how LAWs could reshape societal understandings of warfare and morality. If these weapons become commonplace, there is a risk that the ability of machines to make life-and-death decisions will normalise violence, eroding established norms that protect civilians by maintaining distinctions between combatants and non-combatants. Governments, researchers, and the robotics community must carefully consider these potential impacts as automated warfare becomes more feasible.

The literature review will explore multidimensional arguments that carefully examine the benefits and risks of autonomous weapons systems. The proposed benefits include creating a safer and more controlled environment in war zones and reducing post-traumatic stress disorder (PTSD) among soldiers. Conversely, the risks discussed focus on ethical concerns, an increased likelihood of civilian casualties, and a lack of accountability for actions taken by LAWs. Advocates suggest that LAWs could reduce military personnel's exposure to risk and minimise harm to civilians due to precision targeting and the removal of human error (Burri, 2018, p.160). However, these claims face substantial criticism as by design they are inherently lethal and optimised for military effectiveness. As the degree of human control diminishes, warfare may become dangerously unpredictable and overly reliant on AI-driven decisions which lack the ethical discernment inherent to human operators. Furthermore, present evidence suggests that drones and autonomous 'killer robots' present serious ethical and security challenges (Batabyal, 2024, p.190). The potential for LAWs to make lethal decisions independently raises complex concerns regarding accountability and human rights (Sparrow, 2007). This prospect challenges the assumption that LAWs would provide

net benefits for both combatants and civilians. The following examination of the evidence concludes that whilst automated warfare technology promises some strategic advantages, the risks to established legal and moral norms outweigh these benefits. Without careful ethical oversight, the widespread use of these weapons threatens to redefine warfare, potentially diminishing respect for human life and compromising the longstanding ethical standards that govern armed conflict.

EVALUATING THE SUCCESSES AND FAILURES OF LETHAL AUTONOMOUS WEAPONS

Modern computerised weapon systems can already engage targets with precision beyond human capabilities (United States Navy, 2017). Lee (2018) argues that LAWs equipped with advanced targeting software could perform routine military tasks with far greater accuracy, potentially making warfare more ethical due to a reduced risk of error. In contrast, human soldiers may sometimes act unpredictably or irrationally, influenced by personal judgment and emotions (Umbrello, 2020). Unlike humans, LAWs have the benefit of being unaffected by the emotional and physical toll of war. Given that the main goal of current roboticists is to create weapons which can fully operate independently, Müller (2016) argues that issues like PTSD, trauma, and fatigue (common among human soldiers) would be irrelevant to LAWs. Studies on PTSD and trauma reveal that killing from greater distances reduces the psychological toll associated with killing. Although individual differences may influence this effect, the likelihood of developing psychopathology decreases when soldiers engage in remote killing (Chapa, 2017). This is a prominent argument as researchers conducting interviews with pilots have uncovered that LAWs operators often compare flying drones on the battlefield to playing video games (Bergman, 2016). Currently, however, most LAWs still require some form of human input, and it is therefore unknown whether employing fully independent weapons will resolve psychological disorders associated with war experiences (Phelps, 2021).

Furthermore, research on human behaviour assesses how human soldiers are influenced by various biases. These include stress, fatigue, fear, and personal predispositions which may lead to irrational and unpredictable outcomes during conflicts (Shephard, 2001, p.189). If exposed to prolonged periods of combat, soldiers may use extreme force as a result of excess adrenaline or they may be more sensitive to threat perception due to trauma and stress. This impairs their ability to assess their situation with a clear mind (Shephard, 2001). Therefore, advocates argue that the use of LAWs could reduce instances of impulsive or fear-driven errors on the battlefield as they are unaffected by these cognitive failings. Asaro (2021) supports this view as he argues that technical advantages would render LAWs a safer option, believing that the number of human casualties would be reduced because AI weapons are more controlled and precise.

On the other hand, compelling counterarguments against LAWs focus on ethical and social concerns, particularly the lower threshold for initiating wars. This lower threshold stems from the reduced risks to the attacking side, including the absence of physical harm to personnel and the diminished psychological toll of combat when using LAWs. Horowitz (2021) argues that LAWs also introduce significant uncertainty on the battlefield. Current robotic systems cannot adapt to scenarios outside of their intended design parameters, such as unexpected environmental hazards, degraded communication, or cyber-attacks (van Benthem, 2022). When these systems fail or make errors this can produce unintended consequences, thus raising uncertainty regarding international stability. Such unpredictability lowers the control nations have over managing conflict escalation. Additionally, opponents of LAWs emphasise that the legality of killing in conflict (Solis, 2021) depends on adherence to International Humanitarian Law (IHL). IHL mandates that combatants verify legitimate targets and be held accountable for their actions (Davison, 2018). While LAWs are designed to identify combatants precisely, the complexities of modern battlefields make this task highly challenging. For example, civilians may wear military-style clothing or carry weapons for protection, and likewise, combatants may disguise themselves as civilians (Van Engeland, 2011). These challenges significantly increase the risk of misidentifying targets, raising critical concerns about civilian safety, ethical accountability, and compliance with IHL.

The aforementioned scenarios, which are not uncommon during wars, pose serious challenges for roboticists who favour the deployment of such weapons. While it is anticipated that this technology will continue to evolve and become more intelligent and complex, the likelihood of a weapon being programmed to the extent that it could override its software architecture remains low (Roff, 2014). However, if this were to occur in the future, the machine could potentially disobey orders and independently select its targets in a way similar to human soldiers. This raises a serious concern: if LAWs were able to act independently, they could introduce unpredictable risks, challenging the very principles of control and accountability in warfare. Lin (2008) proposes that in this case, LAWs would simply act like present-day soldiers. However, LAWs could become more dangerous than human soldiers as they are programmed to take down targets, unlike soldiers who have the capability of choosing between causing harm or death and can assess a situation and possibly prevent escalation of conflict. This argument is supported by Sparrow (2007, p.67) who points out that drones and robots do not possess moral judgment and this leaves a lack of interpersonal relationships when attacking a target. Because LAWs lack this relationship with the victims, IHL stipulates that killing by a machine overrules liability and the capability to select legitimate targets. While Sparrow (2007) posits that killing is always immoral, killing without human input further diminishes the victim's dignity.

THE RESPONSIBILITY GAP

Like all weapons in warfare, LAWs are required to comply with IHL. If these systems are unable to meet IHL standards, their use is deemed unlawful and any actions they take could be classified as war crimes (Müller, 2016). Altmann and Sauer (2017) argue that deploying these weapons poses no greater risk to soldiers or civilians than conventional military tactics. However, if LAWs commit war crimes, there remains ambiguity about who would be held accountable. Proponents, such as Arkin (2008), suggest this concern may be unfounded as they envision a future where such technologies outperform humans in identifying legitimate targets and executing commands without the risk of human error.

Nevertheless, the current concern of IHL centres around the lack of responsibility and accountability when killing with these armaments. Weapons such as aerial vehicles can remain in the air for over 24 hours in a war field while being operated by pilots from a computer, at a safe distance from the war zone (Hashimy and Benjamin, 2023, p.2947). This raises the question of who should be held responsible for the atrocities of war when those atrocities are caused by a machine. To date, IHL has not addressed the responsibility gap of who must be held responsible if a LAW wrongly kills a target. Additionally, the United Nations has not reached an agreement on what qualifies as a lawful target for such weapons. This ambiguity highlights a complex moral question: can AI systems apply ethical principles in a meaningful way or make judgments about human life with a moral authority comparable to humans? Some academics argue that only artificial minds could potentially bring true objectivity to ethics (Müller, 2016). Yet, this unresolved issue has contributed to the absence of an international consensus on LAWs, as deciding on whether to deploy such technologies is exceedingly challenging.

Considering the laws of armed conflict, one of the prerequisites is that an agent must be held accountable for a civilian death. Using LAWs without the presence of soldiers further fuels psychological detachment, as the action of pushing a button to commit murder is not as emotionally challenging as fighting face-to-face. As such, ethical apprehensions concerning humanitarian rights arise as military personnel could start to be more emotionally distant from the enemy, potentially creating conditions that facilitate unjustified killings. A study by Singer (2009, p.238) highlights this, in a quote from an interview with a pilot describing his experiences of operating drones from a distance in Iraq: ‘‘It’s like a video game. It can get a little bloodthirsty. But it’s fucking cool”. Research has consistently shown that moral disengagement enhances the probability of acting unethically, as guilt and self-consciousness are diminished (Roykkers, 2010). For example, McAllister (2006) found that individuals who placed less importance on moral arguments were more likely to endorse military attacks against Iraq. Similarly, Aquino et al. (2007) observed that moral disengagement was linked to the 9/11 perpetrators' decision to prioritise death over non-lethal alternatives.

The aforementioned studies suggest that remote killing influences the soldier’s psyche by reducing identification with the victim. Thus, as remote combat is a common feature of today’s warfare, opponents are now regularly perceived as non-human, facilitating the act of killing. However, the full extent to which LAWs aggravate this psychological detachment (compared to more traditional weapons) remains a question that calls for further investigation (Riesen, 2022). If remote weapons were continuously deployed this could lead to a depersonalisation of war, as human soldiers could experience a diminished sense of remorse and accountability. This would be similar to the psychological detachment already observed in operators of LAWs.

THE IMPACT OF LETHAL AUTONOMOUS WEAPONS ON SOCIETY

The deployment of autonomous and remote weapons systems has severe implications for civilian populations, particularly in regions enduring protracted conflict. In areas such as Yemen, Afghanistan, and Pakistan, where U.S. drone warfare has targeted Taliban forces, the persistent presence of drones has fostered profound psychological trauma among civilians (Akbar, 2015). Long-term exposure to drone operations has led to anticipatory anxiety and PTSD across these communities. Civilians describe prolonged periods of stress and fear under the ominous hovering of drones, which has often been associated with unpredictable attacks (Renčelj, 2018). Such exposure has resulted in acute recurrent physiological and psychological symptoms, including fainting, panic attacks, insomnia, hallucinations, and persistent nightmares about seeing drones (Renčelj, 2018).

For those who have suffered personal injury or loss due to drone strikes, these psychological effects are even more severe. These individuals frequently report overwhelming fears that confine them to their homes, inhibiting their ability to participate in community life (Renčelj, 2018). Disturbingly, health professionals in Pakistan and Yemen estimate that over 50% of civilians exhibit symptoms of PTSD, highlighting the widespread mental health crisis exacerbated by drone warfare (Renčelj, 2018). In addition to psychological harm, drone warfare has caused extensive civilian casualties, contradicting claims that autonomous weapons reduce non-combatant fatalities. This is particularly evident in Gaza, where civilian casualties are higher following drone attacks (Heszlein-Lossius et al., 2019). Moreover, the severity of grievous injuries suffered from drone attacks often

surpasses those caused by conventional military methods like tank explosions or soldier attacks (Heszlein-Lossius et al., 2019, p.310).

This evidence challenges the narrative that LAWs decrease collateral damage. Despite advanced targeting algorithms, these systems are programmed to eliminate a target, rather than merely inflict harm. As a result, the force employed by these weapons is far greater than traditional armaments, leading to indiscriminate harm towards both soldiers and civilians (Righetti et al., 2018). It is critical to recognise, however, that research on the mental health impacts of LAWs is still developing and existing studies are often based on small sample sizes. Whilst this research is still valid, the limited samples may reflect the experiences of those in conflict zones who are subjected to constant drone surveillance, rather than those affected by isolated incidents. Supporters of remote weaponry could use this to argue that the military advantages still outweigh the reported harms. However, the lack of research on civilians stems from fear of retribution from oppressive militant movements, which often prevents victims from giving information to researchers (Soeters and Johnson, 2012). This was evident in community members in Afghanistan who voiced their fear of being microchipped by terrorist groups or being beheaded for being informants (Soeters and Johnson, 2012). Consequently, conducting further studies on the mental health impacts of drone warfare risks amplifying the distress of those affected.

Ultimately, proponents of autonomous warfare often overlook the social and psychological toll these weapons exert on civilian communities (Arkin, 2015, p.47). Psychology as a field must contribute to the ethical debate surrounding remote weaponry, considering not only the immediate physical consequences but also the long-term psychological and societal impacts. If the use of advanced weaponry in conflicts continues, it could heighten mistrust between governments and society while exacerbating psychological distress in communities affected by warfare.

CONCLUSION

In conclusion, this review offers a comprehensive examination of arguments for and against the use of autonomous weapons in combat. Benefits of this type of weapon include improved accuracy in detecting targets and elimination of human bias. Furthermore, LAWs have also been linked to psychological benefits for soldiers including a decrease in trauma and a lower mental burden, as employing remote intelligence enables them to be away from areas of conflict. Nonetheless, these proposed benefits are compellingly opposed by studies in both robotics and psychology. As shown by the majority of the present literature, LAWs do not possess the capability to distinguish soldiers who may be disguised as civilians and vice versa, rebutting the claim that such weapons have increased accuracy. Secondly, research on drones emphasises the negative impacts of deploying LAWs, which worsens the physiological and psychological well-being of affected civilians. Thirdly, no government has addressed who must be held responsible if LAWs kill erroneously or do not abide by IHL. Lastly, remote killing contributes to psychological detachment, which leads to a less controlled environment and represents a risk to our moral standards in warfare. Crucially, the impact of such technologies represents a significant ethical risk, extending beyond soldiers to society as a whole, by conveying the message that empathy and compassion have no place in warfare. Ultimately, the widespread deployment of LAWs represents a precarious and uncertain trajectory for the future, one in which we risk sacrificing fundamental human values in the context of conflict.

REFERENCES

Akbar, M. 2015. Drone attacks and suicide bombings: reflections on Pakistan's victims. In: Argomaniz, J. and Lynch, O. eds. *International perspectives on terrorist victimisation: An interdisciplinary approach*. London: Palgrave Macmillan UK, pp.201–224.

Altmann, J. and Sauer, F. 2017. Autonomous weapon systems and strategic stability. *Survival*. **59**(5), pp.117-142.

Aquino, K., Reed, A. II, Thau, S. and Freeman, D. 2007**.** A grotesque and dark beauty: how moral identity and mechanisms of moral disengagement influence cognitive and emotional reactions to war. *Journal of Experimental Social Psychology*. **43**(3), pp.385–392.

Arkin, R. 2009. *Governing lethal behaviour in autonomous robots*. Boca Raton: Chapman and Hall/CRC.

Arkin, R. 2015. The case for banning killer robots: counterpoint. *Communications of the ACM*. **58**(12), pp.46-47.

Asaro, P. 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.

Batabyal, G.S. 2024. Ethics, laws on war and artificial intelligence-driven warfare. In: Roy, K. ed. *Artificial intelligence, ethics and the future of warfare*, Routledge: India, pp.187-215.

Bergman, D. 2016. *The future of drone use: opportunities and threats from ethical and legal perspectives*. Cham: Springer.

Burri, S. 2018. What is the moral problem with killer robots? In: Strawser, J., Jenkins, R. and Robillard, M. eds. *Who should die? The ethics of killing in war*, Oxford: Oxford University Press, pp.163-185.

Chapa, J.O. 2017. Remotely piloted aircraft, risk, and killing as sacrifice: the cost of remote warfare. *Journal of Military Ethics*. **16**(3-4), pp.256-271.

Davison, N. 2018**.** A legal perspective: Autonomous weapon systems under international humanitarian law. In: Nakamitsu, I., Singh Gill, A. and Davison, N. eds. *Perspectives on lethal autonomous weapon systems*. New York: United Nations Office for Disarmament Affairs, pp.360-380.

Gabriel Wood, N. 2020. The problem with killer robots. *Journal of Military Ethics*. **19**(3), pp.220-240.

Hashimy, S.Q. and Benjamin, M.S. 2023. The deployment of US drones in Afghanistan: Deadly sky and unmanned injustice. *International Journal of Law, Management & Humanities*. **6**(2), pp.2942-2958.

Heszlein-Lossius, H., Al-Borno, Y., Shaqqoura, S., Skaik, N., Giil, L.M. and Gilbert, M.F. 2019. Traumatic amputations caused by drone attacks in the local population in Gaza: a retrospective cross-sectional study. *The Lancet Planetary Health*. **3**(1), pp.40-47.

Hijazi, A., Ferguson, C.J., Richard Ferraro, F., Hall, H., Hovee, M. and Wilcox, S. 2019**.** Psychological dimensions of drone warfare. *Current Psychology*. **38**, pp.1285-1296.

Horowitz, M.C. 2021. When speed kills: Lethal autonomous weapon systems, deterrence and stability. In: *Emerging technologies and international stability*. London: Routledge, pp.144-168.

Kormushev, P., Calinon, S. and Caldwell, D.G. 2013. Reinforcement learning in robotics: Applications and real-world challenges. *Robotics*. **2**(3), pp.122-148.

Lee, P. 2018. The distance paradox: Reaper, the human dimension of remote warfare, and future challenges for the RAF. *Air and Space Power Review*. **21**(3), pp.106-130.

Lin, P., Bekey, G. and Abney, K., 2008**.** *Autonomous military robotics: Risk, ethics, and design*. San Luis Obispo: California Polytechnic State University.

McAllister, I. 2006. A war too far? Bush, Iraq, and the 2004 US presidential election. *Presidential Studies Quarterly*. **36**(2), pp.260-280.

Michalec, O., O'Donovan, C. and Sobhani, M. 2021. What is robotics made of? The interdisciplinary politics of robotics research. *Humanities and Social Sciences Communications*. **8**(1).

Müller, V.C. 2016. Autonomous killer robots are probably good news. In: Di Nucci, E. and de Sio, F. S. eds. *Drones and responsibility*. London: Routledge, pp.67-81.

Renčelj, M. 2018. *Permanent presence of armed drones and the elusive concept of psychological harm*. Canada. McGill University.

Riesen, E. 2022. The moral case for the development and use of autonomous weapon systems. *Journal of Military Ethics*. **21**(2), pp.132-150.

Righetti, L., Pham, Q.C., Madhavan, R. and Chatila, R. 2018. Lethal autonomous weapon systems: ethical, legal, and societal issues. *IEEE Robotics & Automation Magazine*. **25**(1), pp.123-126.

Roff, H.M. 2014. The strategic robot problem: lethal autonomous weapons in war. *Journal of Military Ethics*. **13**(3), pp.211-227.

Royakkers, L. and Van Est, R. 2010. The cubicle warrior: the marionette of digitalized warfare. *Ethics and Information Technology*. **12**, pp.289-296.

Shephard, B. 2001. *A war of nerves: soldiers and psychiatrists in the twentieth century*. Cambridge: Harvard University Press.

Singer, P.W. 2009. *Wired for war: the robotics revolution and conflict in the 21st century.* London: Penguin.

Soeters, J. and Johnson, H.N. 2012. Informerphobia: understanding the 'fear factor' in crime-and terrorism-related information disclosure in Afghanistan. In: Kümmel, G. and Soeters, J. eds. *New wars, new militaries, new soldiers: conflicts, the armed forces and the soldierly subject.* Basingstoke: Palgrave Macmillan, pp.167-184.

Solis, G.D. 2021. *The law of armed conflict: international humanitarian law in war*. Cambridge: University Press.

Sparrow, R. 2007. Killer robots. *Journal of Applied Philosophy*. **24**(1), pp.62-77.

Umbrello, S., Torres, P. and De Bellis, A.F. 2020. The future of war: could lethal autonomous weapons make conflict more ethical? *AI & Society*. **35**(1), pp.273-282.

United States Navy, 2017. Autonomous weapons are already here. *World Economic Forum*. [Online]. [Accessed 21 February 2025]. Available at: https://www.weforum.org/stories/2017/11/autonomous-weapons-are-already-here-but-humans-are-still-needed-on-the-battlefield/

van Benthem, T.J. 2022. Exploring changing battlefields: autonomous weapons, unintended engagements and the law of armed conflict. In: Jančárková, T., Visky, G. and Winther, I. eds. *14th international conference on cyber conflict: keep moving! (CyCon), 31 May – 3 June 2022, Tallinn*. pp.190-204.

Van Engeland, A. 2011. *Civilian or combatant?: A challenge for the 21st century*. Oxford: Oxford University Press.

Winter, E. 2022. The compatibility of autonomous weapons with the principles of international humanitarian law. *Journal of Conflict and Security Law*. **27**(1), pp.1-20.