

Truth, Lies and New Weapons Technologies: Prospects for *Jus in Silico*?

Studies in Christian Ethics

1–19

© The Author(s) 2021



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/09539468211051240

journals.sagepub.com/home/sce**Esther D. Reed** 

Theology and Religion, University of Exeter, Exeter, UK

Abstract

This article tests the proposition that new weapons technology requires Christian ethics to dispense with the just war tradition (JWT) and argues for its development rather than dissolution. Those working in the JWT should be under no illusions, however, that new weapons technologies could (or do already) represent threats to the doing of justice in the theatre of war. These threats include weapons systems that deliver indiscriminate, disproportionate or otherwise unjust outcomes, or that are operated within (quasi-)legal frameworks marked by accountability gaps. The temptation to abrogate (L. *abrogare*—repeal, evade) responsibility to the machine is also a moral threat to the doing of justice in the theatre of war.

Keywords

Just war, justice, new weapons technologies, automation, accountability, anthropomorphism

Introduction

Robots, machine-learning, and digital technologies more generally, are bringing challenges to the weapons industry and the military, policy-makers and governments.¹

1. 'Whilst it is envisaged that humans will continue to be central to the decision-making process', observes the UK's Development, Concepts and Doctrine Centre, 'conflicts fought increasingly by robots or autonomous systems could change the very nature of warfare'. Ministry of Defence, 'Global Strategic Trends: The Future Starts Today', 6th edn, October 2018, p. 18, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/771309/Global_Strategic_Trends_-_The_Future_Starts_Today.pdf (accessed 20 August 2021).

Corresponding author:

Esther D. Reed, Theology and Religion, University of Exeter, Amory Building, Exeter EX4 4QJ, UK.

Email: e.d.reed@exeter.ac.uk

Levels of automation in weapons systems (will) vary,² but weapons systems with artificial intelligence (AI) capability, operated autonomously to find and attack humans, have probably already been deployed.³ In March 2021, a United Nations report from the panel of experts on Libya indicated that an AI drone that attacked fighters may have acted on its own. To cite directly from this report: ‘The lethal autonomous weapons systems were programmed to attack targets without requiring data connectivity between the operator and the munition: in effect, a true “fire, forget and find” capability, i.e. that, once fired, is able to guide itself to its target’.⁴ This report does not say more about the capability or whether there were civilian casualties or injuries, but it appears that the machine selected targets and decided to kill.

Existing International Humanitarian Law (IHL) holds state or other military actors responsible for whatever is done with any weapon but does not contain any explicit prohibition against lethal autonomous weapons or AI-enabled weapons systems. Article 36 under the Protocol Additional to the Geneva Conventions of 12 August 1949 requires a High Contracting Party to determine whether the employment of any new weapon would, in some or all circumstances, be prohibited by any applicable rule of international law. But, of course, existing IHL was not developed or tested for weapons technology whereby weapons ‘after initial activation, select and apply force to targets without human intervention, in the sense that they are triggered by their environment based on a “target profile”, which serves as a generalised approximation of a type of target’.⁵ Yet future military success is likely to depend on operational advantage achieved through human-machine interaction and unmanned weapons.

Thirty countries are calling for a ban on fully autonomous weapons and for the creation of a new ban treaty to establish the principle of meaningful human control over

-
2. As Yoram Dinstein and Arne Willy Dahl acknowledge in the *Oslo Manual on Select Topics of the Law of Armed Conflict: Rules and Commentary*, ‘The technologies ... are either novel or have yet to emerge and not explicitly dealt with in either customary international law or in treaty law’. Y. Dinstein and A.W. Dahl, *Oslo Manual on Select Topics of the Law of Armed Conflict: Rules and Commentary*, 2020, p. 31, https://library.oapen.org/bitstream/id/7065a5e9-c8a8-4cc7-a7f9-7f8f4fadd056/2020_Book_OsloManualOnSelectTopicsOfTheL.pdf (accessed 20 August 2021).
 3. M. Cramer, ‘A.I. Drone May Have Acted on Its Own in Attacking Fighters, U.N. Says’, *The New York Times*, 3 June 2021, <https://www.nytimes.com/2021/06/03/world/africa/libya-drone.html> (accessed 20 August 2021).
 4. Letter dated 8 March 2021 from the Panel of Experts on Libya established pursuant to resolution 1973 (2011) addressed to the President of the Security Council S/2021/229. HAF = Haftar Armed Forces, i.e., part of the Libyan National Arab Army under the command of Field Marshal Khalifa Haftar. <https://digitallibrary.un.org/record/3905159?ln=en> (accessed 20 August 2021).
 5. International Committee of the Red Cross, ‘Autonomous Weapons: The ICRC Recommends Adopting New Rules’, Statement, Switzerland, 3 August 2021, <https://www.icrc.org/en/document/autonomous-weapons-icrc-recommends-new-rules> (accessed 20 August 2021).

the use of force.^{6,7} This list is notable, however, for who have declined.⁸ The United States, the United Kingdom, Israel, Australia and Russia are opposed to a pre-emptive ban. China's position is less clear; they have called for a ban on use and not on the development of these weapons.⁹ The UK's stated position towards fully autonomous weapon systems is that 'the operation of our weapon systems will always be under human control and no UK weapons systems will be capable of attacking targets without this'.¹⁰

-
6. Campaign to Stop Killer Robots, 'The Threat of Fully Autonomous Weapons', 2021, <https://www.stopkillerrobots.org/learn/#problem> (accessed 20 August 2021).
 7. The phrase 'meaningful human control' is significant. On 4 February 2021, the World Council of Churches opened an interfaith statement for signature entitled 'A Plea for Preserving Our Shared Humanity' which called for all people of goodwill 'to commit to preserving meaningful human control over the use of force, and enact a pre-emptive ban on fully autonomous weapons'. See <https://www.oikoumene.org/sites/default/files/2021-02/Interfaith%20Statement%20on%20Killer%20Robots%20%28February%202021%29.pdf> (accessed 20 August 2021). The phrase has been used for much longer in relation to lethal autonomous weapons systems (LAWS) by the UK-based NGO Article 36, which is pressing for international regulation that prohibits autonomy and weapons systems 'where target profiles are built on the basis of machine learning such that commanders do not know the actual characteristics of conditions under which they will apply force (i.e. they are not able to assess the presence of conditions that will produce false positives)'. See R. Moyes, Article 36 Policy Note, 'Autonomy in Weapons Systems: Mapping a Structure for Regulation through Specific Policy Questions', November 2019, <https://article36.org/wp-content/uploads/2019/11/regulation-structure.pdf> (accessed 20 August 2021). Article 36 is similarly concerned about 'systems where targeted profiles, the conditions under which force will be applied, change within the system after it has been put to use and not approved by human commander'. Ibid. See also International Committee of the Red Cross, 'Autonomous Weapons: States Must Agree on What Human Control Means in Practice', 20 November 2018, <https://www.icrc.org/en/document/autonomous-weapons-states-must-agree-what-human-control-means-practice> (accessed 20 August 2021).
 8. The thirty countries calling for a ban on killer robots are: Algeria, Argentina, Austria, Bolivia, Brazil, Chile, China (use only), Colombia, Costa Rica, Cuba, Djibouti, Ecuador, El Salvador, Egypt, Ghana, Guatemala, the Holy See, Iraq, Jordan, Mexico, Morocco, Namibia, Nicaragua, Pakistan, Panama, Peru, the State of Palestine, Uganda, Venezuela, and Zimbabwe. Human Rights Watch, 'Killer Robots: Growing Support for a Ban', 10 August 2020, <https://www.hrw.org/news/2020/08/10/killer-robots-growing-support-ban> (accessed 20 August 2021).
 9. According to Melissa K. Chan, 'Its delegates contribute just the minimum, and often send ambiguous signals on where they stand. They have called killer robots a "humanitarian concern", yet have stepped in to water down the text being debated'. M. K. Chan, 'China and the U.S. are Fighting a Major Battle over Killer Robots and the Future of AI', *Time*, 13 September 2019, <https://time.com/5673240/china-killer-robots-weapons/> (accessed 20 August 2021).
 10. Ministry of Defence Joint Doctrine Publication 0-30.2, 'Unmanned Aircraft Systems', August 2017, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/673940/doctrine_uk_uas_jdp_0_30_2.pdf (accessed 24 August 2021); Ministry of Defence Letter, 4 January 2021, <https://article36.org/wp-content/uploads/2021/01/UK-govt-reply-2020-LAWS.pdf> (accessed 24 August 2021). See also Government of

The UK Ministry of Defence's position is as follows:

It is our view that we should embrace and welcome technological advancements that can support compliance with International Humanitarian Law, as in doing so we can offer greater protection to persons who are not, or who are no longer, participating in hostilities. A legally binding instrument which hampers the legitimate development and use of such technologies would be counterproductive.¹¹

A plausibly universal ban has not yet been achieved within the United Nations framework via the CCW. Hence the derision heaped upon the High Contracting Parties to the CCW by non-profit organizations such as the Arms Control Association, Article 36, and substantial political pressure via other organizations.¹² In a significant move on 3 August 2021, however, the International Committee of the Red Cross (ICRC) recommended that states adopt new, legally binding rules to regulate autonomous weapon systems to ensure

the United Kingdom, 'Statement to the Convention on Conventional Weapons informal meeting of experts on lethal autonomous weapons systems', 15 November, 2017, https://conf.unog.ch/digitalrecordings/index.html?guid=public/61.0500/EDA01C29-125B-4280-9EC1-92C44A4E9830_10h09&position=3223 (accessed 24 August 2021); Government of the United Kingdom, 'Statement to the Convention on Conventional Weapons' informal meeting of experts on lethal autonomous weapons systems, 13 April 2015, http://www.reachingcriticalwill.org/images/documents/Disarmament-fora/ccw/2015/meeting-experts-laws/statements/15April_UK.pdf (accessed 16 June 2020).

11. Ministry of Defence Letter, 4 January 2021, <https://article36.org/wp-content/uploads/2021/01/UK-govt-reply-2020-LAWS.pdf>. For a good summary, see H. Evans, 'Too Early for a Ban: The U.S. and U.K. Positions on Lethal Autonomous Weapons Systems', *Lawfare*, 13 April 2018, <https://www.lawfareblog.com/too-early-ban-us-and-uk-positions-lethal-autonomous-weapons-systems> (accessed 24 August 2021).
12. The Arms Control Association reported that some parties, as well as civil society representatives, derided the lack of progress toward a more comprehensive prohibition: 'States parties to the Convention on Certain Conventional Weapons (CCW) discussed in November [2019] the danger of lethal autonomous weapons systems, often referred to as "killer robots", but could not agree to begin a formal process of creating a legal instrument to govern their development and use. This is the fifth year in which the topic has been addressed by the CCW, which seeks to control especially harmful or indiscriminate conventional arms'. Arms Control Association, 'Decision on Autonomous Weapons Talks Eludes CCW', 2019, <https://www.armscontrol.org/act/2019-12/news-briefs/decision-autonomous-weapons-talks-eludes-ccw> (accessed 16 June 2020). In his 'Critical Commentary on the "Guiding Principles"', Director of Article 36 Richard Moyes writes: 'That assertions like this can become a comfortable mantra for states is an example of the erosive capacity of the CCW and similar multilateral frameworks. No state actually argued in the informal discussions or the GGE [Group of Government Experts] that IHL did not apply to weapon technologies in this area. Quite the opposite, many noted that it did—perhaps simply relieved to have something straightforward and incontestable to write into their own statements'. R. Moyes, Article 36 Policy Note, 'Critical Commentary on the "Guiding Principles"', November 2019, <http://www.article36.org/wp-content/uploads/2019/11/Commentary-on-the-guiding-principles.pdf> (accessed 17 July 2021).

that sufficient human control and judgement is retained in the use of force: 'It is the ICRC's view that this will require prohibiting certain types of autonomous weapon systems and strictly regulating all others'.¹³ Religious voices, notably the Pax Christi peace movement, have campaigned for well over a decade alongside the 'Stop Killer Robots' campaign for a pre-emptive ban, for a range of ethical, legal and security reasons.¹⁴ The Vatican warns that the delegation of powers to autonomous systems 'puts us on the path of negation, oblivion, and contempt for the essential characteristics unique to the human persons and to soldierly virtues'.¹⁵ In November 2018, Archbishop Ivan Jurkovic called for a ban on the development of robotic weapons which have 'the capacity of altering irreversibly the nature of warfare, becoming more detached from human agency, putting in question the humanity of our societies'.¹⁶ But on the other side, counter arguments have been voiced by those who emphasize that potential adversaries are far advanced in their development of these technologies, that there is no such thing as a fully autonomous system because a human is always in the loop somewhere, and that the necessary and appropriate ethical focus is on the human role in authorizing lethal force and controlling it. Debate focuses on whether an outright ban is feasible, whether new legal rules are needed or the weapons in question can be operated under existing IHL requirements because what matters is accountability and the weapons-control system as already the means by which human control is maintained.¹⁷

Christian people, and global citizens more generally, have three possible routes ahead: (i) lending support to political pressure for an outright ban on the development and use of LAWS in the hope that all countries of the world, including the major military powers, become signatories; (ii) lending support to political pressure for a legally binding international treaty agreed by many countries of the world, including the major military powers, that sets out agreed parameters within which autonomous weapons can be developed and used; (iii) accepting that political stalemate leads to 'a non-binding political statement or to a position of permanent stasis'.¹⁸ How closely the practices of Christian pacifism and just war reasoning map on to options (i) and (ii) is an open question, but both will surely keep company for good part of the way in striving for an outright

13. International Committee of the Red Cross, 'Autonomous Weapons'.

14. Pax for Peace, 'Stop Killer Robots?', <https://paxforpeace.nl/what-we-do/programmes/killer-robots> (accessed 24 August 2021).

15. J. Fair, 'Archbishop Jurkovic Warns of Autonomous Weapons', 11 April 2018, <https://zenit.org/articles/archbishop-jurkovic-warns-of-autonomous-weapons/> (accessed 24 August 2021).

16. Catholic News Service, 'Vatican Official: Prohibit "Killer Robots" Now Before They Become Reality', *The Tablet*, 28 November 2018, <https://www.thetablet.co.uk/news/11072/vatican-official-prohibit-killer-robots-now-before-they-become-reality> (accessed 24 August 2021).

17. T. McFarland and J. Galliot, 'Understanding AI and Autonomy: Problematising the Meaningful Human Control Argument against Killer Robots', in J. Galliot, D. MacIntosh and J.D. Ohlin (eds.), *Lethal Autonomous Weapons: Re-Examining the Law and Ethics of Robotic Warfare* (Oxford: Oxford University Press, 2021), p. 51.

18. These are basically the three options laid out at <https://www.e-ir.info/2020/04/15/introducing-guiding-principles-for-the-development-and-use-of-lethal-autonomous-weapon-systems/> (accessed 24 August 2021).

ban.¹⁹ Option (iii) would, I suggest, be unacceptably neglectful of the expectation upon all Christian people to be peacemakers and proactive in weapons control and matters pertaining to control over weapons; persistence is required (Lk. 18:1-8).²⁰

Pope Francis wrote in *Fratelli Tutti* of his concerns about the development of nuclear, chemical and biological weapons, and the enormous and growing possibilities offered by new technologies such that, in the words of Pope John XXIII, 'it no longer makes sense to maintain that war is a fit instrument with which to repair the violation of justice'.²¹ In this article, I do not consider nuclear, chemical and biological weapons, the inherently indiscriminate nature of which makes limited war more-or-less impossible, but concentrate on LAWS. At issue is whether options (i) and (ii) above are available to Christian people, or only option (i). The following questions are addressed:

1. Do new weapons technologies that include some level of automation make it prohibitively difficult to maintain the just war tradition (JWT) as a proposal for doing justice in the theatre of war?
2. What about this weaponry is peculiarly unethical?
3. Can/should the JWT evolve as *jus in silico*? If so, how?
4. What is the ministry of the Church and the peculiar responsibilities of Christian ethics amidst these challenges?

At its broadest, the position developed is that the JWT in Christian perspective remains a proposal for doing justice in the theatre of war and is needed for ethical consideration of the prohibition and regulation of non-prohibited LAWS. Those working in the JWT should be under no illusions, however, that new weapons technologies could (or do already) represent threats to the doing of justice in the theatre of war. These threats include weapons systems that deliver indiscriminate, disproportionate or otherwise unjust outcomes, or that are operated within (quasi-)legal frameworks marked by accountability gaps. The temptation to abrogate (L. *abrogare*—repeal, evade) responsibility to the machine is also a moral threat to the doing of justice in the theatre of war.

Key Ethical and Legal Challenges Posed by New Technologies

To address the question of whether new weapons technologies that include some level of automation make it prohibitively difficult to maintain the JWT as a proposal for doing justice in the theatre of war, it is necessary to consider key ethical and legal challenges posed by new technologies. The challenges resulting from new weapons technologies

19. Karl Barth, *Church Dogmatics* (Edinburgh: T&T Clark, 1961), III/4, §55 'The Protection of Life', p. 458.

20. By weapons control, I broadly understand to mean international restrictions upon the development, production, stockpiling, proliferation and usage of various weapons types. Control over weapons has to do with clear lines of responsibility and accountability for actions performed with weapons, and the capacity to disarm them.

21. Pope Francis, *Fratelli Tutti*, 3 October 2020, para 258, https://www.vatican.va/content/francesco/en/encyclicals/documents/papa-francesco_20201003_enciclica-fratelli-tutti.html (accessed 24 August 2021).

may be broadly categorized into three groups. First are the legally-framed operational issues of the kind that might be addressed by Parties to the CCW for the lawful constraint of weapons capability.²² Second are the global political issues pertaining to how AI-enhanced weapons systems are likely to lower barriers to entering conflict in ways that undermine the international rule of law, especially in an era of ‘persistent competition’ below the threshold of war.²³ Third are personnel-focused issues, that is, those bearing upon the experiences of the serving military. This article is part of a larger project in which I attempt to consider all of the above. In the next section I consider briefly definitional issues wherein anthropomorphism tends towards the abrogation of responsibility (i.e., anthropocentric terminology applied to technology serves to reduce or even eliminate the spaces necessary for human agency, perhaps even to disguise that elimination), before turning to the detachment of war fighting from human agency and related concerns about accountability.

Definitions of Autonomy

All ethical concerns in this area begin to some extent with definition(s), but debate is hindered by the lack of intergovernmental agreement on the definition of autonomous weapons systems. This matters because ethical issues are entailed in definitions. In 2017, the UK Ministry of Defence paper ‘Unmanned Aircraft Systems’ distinguished between automated and autonomous systems:

Automated system

In the unmanned aircraft context, an automated or automatic system is one that, in response to inputs from one or more sensors, is programmed to logically follow a predefined set of rules in order to provide an outcome. Knowing the set of rules under which it is operating means that its output is predictable.

Autonomous system

An autonomous system is capable of understanding higher-level intent and direction. From this understanding and its perception of its environment, such a system is able to take appropriate action to bring about a desired state. It is capable of deciding a course of action, from a number of alternatives, without depending on human oversight and control, although these may still be present. Although the overall activity of an autonomous unmanned aircraft will be predictable, individual actions may not be.²⁴

This distinction was described by the House of Lords Artificial Intelligence Committee report (HL Paper 100) ‘AI in the UK: Ready, Willing and Able?’ as ‘a

22. Fair, ‘Archbishop Jurkovic Warns of Autonomous Weapons’.

23. Phrase used by General Sir Nick Carter, Chief of Defence Staff, RUSI Annual Lecture 2020, <https://www.gov.uk/government/speeches/chief-of-defence-staff-at-rusi-annual-lecture> (accessed 24 August 2021).

24. Ministry of Defence Joint Doctrine Publication 0-30.2, ‘Unmanned Aircraft Systems’.

relatively unusual distinction'.²⁵ The report further listed definitions of autonomous lethal weapon systems used by Austria, France, the Holy See, Italy, the Netherlands, Norway, Switzerland and the USA, and observed: 'none would appear to set the bar as high as the UK. All of these definitions focus on the level of human involvement in supervision and target setting, and do not require "higher level intent and direction", which could be taken to mean at least some level of sentience'.²⁶ In other words, the House of Lords Select Committee was critical of the UK's unusual definition and recommended that the UK definition of autonomous weapons should be realigned to be the same, or similar, as that used by the rest of the world: 'Without agreed definitions we could easily find ourselves stumbling through a semantic haze into dangerous territory'.²⁷

With the UK's definition being situated so far in the technological 'future' (to the point of perhaps being unrealizable, by the MoD's own admission), statements such as 'we have no plans to develop or acquire such weapons' could appear progressive without actually applying any constraint on the UK's ability to develop weapons systems with greater and greater autonomy.²⁸ Indeed, brief comparison with the definition of an autonomous weapons system adopted by the Holy See further makes the point. In a working paper submitted to the United Nations office at Geneva in April 2016, the Holy See defined an autonomous weapon system as 'a weapon system capable of identifying, selecting and triggering action on a target without human supervision'.²⁹ No reference is made to 'higher-level intent and direction' but it is different again from that used, for instance, by Austria: 'Autonomous weapons systems (AWS) are weapons that, in contrast to traditional inert arms, are capable of functioning with a lesser degree of human manipulation and control, or none at all'.³⁰

In 2020, the *Oslo Manual on Select Topics of the Law of Armed Conflict: Rules and Commentary*, authored by Yoram Dinstein and Arne Willy Dahl, offered this definition:

For the purposes of this Manual, an 'autonomous' weapon system is a weapon system that is programmed to apply human-like reasoning to determine whether an object or person is a target, whether it should be attacked, and if so, how and when.³¹

25. Parliamentary Business, 'AI in the UK: Ready, Willing and Able?', 2018, <https://publications.parliament.uk/pa/ld201719/ldselect/ldai/100/10012.htm> (accessed 20 August 2021). See further Box 11: Definitions of lethal autonomous weapons systems used by other countries.

26. Parliamentary Business, 'AI in the UK', para. 342.

27. Parliamentary Business, 'AI in the UK', para. 325.

28. Article 36, 'Shifting Definitions – the UK and Autonomous Weapons Systems', July 2018, <https://article36.org/wp-content/uploads/2018/07/Shifting-definitions-UK-and-autonomous-weapons-July-2018.pdf> (accessed 20 August 2021).

29. Holy See, 'Elements Supporting the Prohibition of Lethal Autonomous Weapons Systems' (Working Paper, 7 April 2016).

30. Parliamentary Business, 'AI in the UK', Box 11. For discussion on the meaning of 'intentionality' and 'higher order intentionality', see Catrin Misselhorn, 'Collective Agency and Co-operation in Natural and Artificial Systems', in C. Misselhorn (ed.), *Collective Agency and Cooperation in Natural and Artificial Systems: Explanation, Implementation and Simulation* (Switzerland: Springer International Publishing, 2015), pp. 6–8.

31. Dinstein and Dahl, *Oslo Manual on Select Topics of the Law of Armed Conflict*, p. 34.

Note here the language of ‘human-like reasoning’, that is, the supposition that machines have internal states that correspond to human reason. Ethical dangers attaching to anthropomorphism in the context of decision-making are considerable. Anthropomorphism which supposes ‘higher-level intent’ or ‘human-like reasoning’ in an ‘autonomous’ weapon system must be challenged not least because it potentially undermines the CCW principle that accountability cannot be transferred to machines and could influence individual and societal attitudes toward machines which surpass human ability in computational problems, data search capacity, surveillance, the transformation of information into usable knowledge, and more.

For present purposes, I adopt the following EU definitions:

‘AI system’ means a system that is either software-based or embedded in hardware devices, and that displays behaviour simulating intelligence by, inter alia, collecting and processing data, analysing and interpreting its environment, and by taking action, with some degree of autonomy, to achieve specific goals;

‘Autonomous’ means an AI system that operates by interpreting certain input, and by using a set of predetermined instructions, without being limited to such instructions, despite the system’s behaviour being constrained by and targeted at fulfilling the goal it was given and other relevant design choices made by its developer.³²

These EU definitions avoid anthropomorphic terms for machine-learning systems in which similarities to human cognition are ‘vastly overstated and narrowly construed’.³³ A primary concern, however, remains the anthropomorphism that tends towards the abrogation of responsibility to the machine, thereby raising questions about the meaning of ‘human agency’ and ‘accountability’.

Anthropomorphism

It is commonplace and almost unavoidable these days to describe everyday AI in terms of human capabilities and traits. Alan Turing’s 1950s ‘imitation game’ (now called the Turing Test) tested a machine’s ability to exhibit intelligent behaviour by whether it was equivalent to, or indistinguishable from, that of a human.³⁴ Today, owners of smartphones ask questions of the cloud-based voice services and hear a natural-sounding voice telling us the weather forecast or traffic conditions ahead, or might watch the Boston

32. European Parliament Plenary Sitting, ‘Report on Artificial Intelligence: Questions of interpretation and application of international law in so far as the EU is affected in the areas of civil and military uses and of state authority outside the scope of criminal justice’, 4 January 2021, https://www.europarl.europa.eu/doceo/document/A-9-2021-0001_EN.pdf (accessed 20 August 2021).

33. Phrase borrowed from David Watson, ‘The Rhetoric and Reality of Anthropomorphism in Artificial Intelligence’, *Minds and Machines* 29 (2019), pp. 417–40, at p. 417.

34. A. Turing, ‘Computing Machinery and Intelligence’, *Mind* LIX.236 (October 1950), pp. 433–60, <https://academic.oup.com/mind/article/LIX/236/433/986238> (accessed 20 August 2021).

Dynamic robots perform dance moves and exclaim with surprise at how the robot moves its 'hands' and 'feet'.³⁵ Casual anthropomorphism is probably unavoidable and relatively harmless. Of concern here is not the understandable tendency of describing how a machine picks up an item or moves with the language of 'hands' and 'feet' but interactions with, and inferences from, computational programs that could make users vulnerable to being steered in particular directions when making decisions.³⁶ It is that machine-learning, simulation, and computing infrastructures described as AI employing mathematical systems that act and think like humans implies that these machines have minds which do things that humans do, only better. Understanding AI through the lens of human mental features risks reducing it to a sort of replica of the human mind and leads to a flawed and ultimately limited ethical analysis of the issues AI raises.³⁷ Anthropomorphism can tend toward the placing of blind faith in the accuracy of systems whose decision-making processes are not available to scrutiny (see 'black box' problem below); the anthropomorphic language used of weapons systems risks masking important limitations intrinsic to machine-learning and making it easier to convince users to trust their capabilities.³⁸

Identifying appropriate descriptors for machine-human interaction in war fighting is a larger topic than can be tackled here. For the moment, my plea is for descriptors that are accurate and that resist the implication that neural-network training or 'deep learning' in weapons systems is a form of 'human-like' reasoning. Getting the language right is one step toward not abrogating responsibility to the machine, creating false expectations around the e-trust relationship that is appropriate with a machine, or reducing ethical reasoning to probability calculus and risk management.³⁹ This is especially important as machines are programmed to make choices recognized as morally good rather than morally bad, and anthropomorphic descriptions of robots as morally good are becoming familiar.⁴⁰ For instance, Duncan MacIntosh, in his discussion of 'fire and forget' weapons systems writes: 'For, given the choice between control by a morally bad human who

35. 'Watch Boston Dynamic robots show off dance moves in new viral video', 30 December 2020, <https://www.youtube.com/watch?v=BFK9lkez32E> (accessed 20 August 2021).

36. This problem is considered in A. Salles, K. Evers and M. Farisco, 'Anthropomorphism in AI', *AJOB Neuroscience* 11.2 (2020), pp. 88–95 (see especially p. 90), <https://www.tandfonline.com/doi/full/10.1080/21507740.2020.1740350?scroll=top&needAccess=true> (accessed 20 August 2021).

37. Salles, Evers and Farisco, 'Anthropomorphism in AI'.

38. D. Proudfoot, 'Anthropomorphism and AI: Turing's Much Misunderstood Imitation Game', *Artificial Intelligence* 175.5–6 (2011), pp. 950–57.

39. See, for example, M. Taddeo, 'Modelling Trust in Artificial Agents, A First Step Toward the Analysis of e-Trust', *Minds and Machines* 20 (2010), pp. 243–57. 'This paper provides a new analysis of e-trust, trust occurring in digital contexts, among the artificial agents of a distributed artificial system. The analysis endorses a non-psychological approach and rests on a Kantian regulative ideal of a rational agent, able to choose the best option for itself, given a specific scenario and a goal to achieve'.

40. A. Sharkey, 'Can We Program or Train Robots to Be Good?', *Ethics and Information Technology* 22 (2020), pp. 283–95.

would kill someone undeserving of being killed and a morally good robot who would kill only someone deserving of being killed, we would pick the good robot'.⁴¹ His contentious point is that compliance with IHL can and should be encoded. Indeed, MacIntosh goes further to argue that it is not always morally better for a human decision to be proximal to the application of force.

In the larger project of which this article is part, the challenge of why and how to denounce the reduction of ethics to technical practice—what Pope Francis calls the 'technocratic paradigm'⁴²—is taken as part of a discussion of why and how ethical reasoning isn't just about applying rules and calculating possible outcomes, as if reducible to an exercise in logic or mathematics, but is about relationships with God and neighbour. That AI already outperforms humans in various tasks of description and prediction, can be programmed to discriminate between objects etc., is potentially of tremendous benefit. But machine-learned ability to acquire and apply knowledge in decision-making is not (yet) the same as having a human conscience and developing the relationships and character traits required more broadly of human decision-makers; the structure of the act of judgement, which includes consideration *inter alia* of end, object, means and circumstance, is about more than the tasks for which machines can be programmed.⁴³

As a first step, it is important to be clear that the so-called 'blackbox problem' where weapons systems are designed such that unpredictability and nonavailability to scrutiny are built-in, is not the same as the problems of human decision-making, for example, a split-second decision about whether or not to fire taken under extreme pressure in urban warfare; the descriptor 'human-like reasoning' used of machine-learning processes can potentially obscure the differences between these decision-making processes. Algorithms allow software to learn from patterns or features in the data. So-called neural-network training or 'deep learning' in weapons systems will have 'learned' from training data how to distinguish between a spade and a rifle, a military tank and a hay cart, or

41. D. MacIntosh, 'Fire and Forget: A Moral Defence of the Use of Autonomous Weapons Systems in War and Peace', in Galliot, MacIntosh and Ohlin, *Lethal Autonomous Weapons*, p. 10. For a discussion of human attitudes to behaviours in robots, see J. Banks, 'Good Robots, Bad Robots: Morally Valenced Behavior Effects on Perceived Mind, Morality, and Trust', *International Journal of Social Robotics* (2020).

42. Pope Francis, *Laudato Si'*, 18 June 2015, https://www.vatican.va/content/francesco/en/encyclicals/documents/papa-francesco_20150524_encyclica-laudato-si.html (accessed 20 August 2021). See also Vatican News, 'Pontifical Academy for Life: Yes to Artificial Intelligence, But with Ethics', 17 May 2020, <https://www.vaticannews.va/en/vatican-city/news/2020-05/artificial-intelligence-with-ethics-pontifical-academy-for-life.html> (accessed 20 August 2021), and Elke Schwarz on 'ethics as technics', i.e., ethics 'rendered as code and facilitated through algorithmic operations' in ways that tend toward putting ethics 'beyond contestation through its reliance on professionalism and ostensibly superior modes of technology'. E. Schwarz, 'Ethics as Technics in Death Machines', 11 July 2018, ch. 5, <https://www.manchesterhive.com/view/9781526114839/9781526114839.00010.xml> (accessed 20 August 2021).

43. Aquinas, *ST II-II*, q.60, <https://www.newadvent.org/summa/3060.htm> (accessed 20 August 2021).

suchlike. The computer might be given 1000 or 100,000, or millions more, pictures of spades and rifles, military tanks and hay carts, until it is capable with a higher level of predictability of identifying the correct object. That a human's decision-making processes might be biased (for good or ill), clouded by fear, anger, revenge, and so on, is a relevant consideration. But how 'human-like' this process is remains open to question. So too is the machine's supposed 'human-like' capacity to deal with unpredictability, novelty in the real-world situation, anomaly in new data, and more. At the least we must say that the inaccessibility of decision-making routes to subsequent analysis, namely, the impossibility of knowing, either before or after the event/output, how the machine processed the data to reach its decision, does not equate to being 'human-like'.

Detachment of War Fighting from Human Agency

If we understand the problem of agency broadly as losing sight of connections between agent, act and consequence, then the problem is not exclusive to the military; linear connections from agent to act to consequence are increasingly untraceable not only in AI but in many features of our era.⁴⁴ In the context of new weapons technologies, the problem of agency has peculiar force because of the demand before God for accountability for the taking of human life (Gen. 9:5-6).⁴⁵ Has a moral line been crossed when the machine processes information and 'decides' to fire? Are machine-executed acts non-agent 'acts' (necessarily) like trees falling in forests, rocks tumbling down hills, computers 'acting' in inscrutable ways detached from human control? Is this moral line such that Christian ethics removes itself from the debate to concentrate only on 'ban killer robots' campaigns? Is the reality more complex such that concentrating only on 'ban killer robots' campaigns would be a different type of derogation of responsibilities?

This cluster of problems faces us with a choice. Two potentially defensible routes are available: (1) the just war tradition *is not* capable of addressing issues raised by LAWS because the technology entails inherently unethical decisions and actions (*mala in se*),⁴⁶ (2) the just war tradition *is* capable of addressing issues raised by LAWS because of decision-making power with respect to choosing the lesser of evils (*minima de malis*) and should do so by striving for guiding principles and international regulation to govern the development and use of this technology. (No useful purpose is served by discussing whether the JWT is capable of evolving without either advocating an outright ban

44. The problem of agency will be familiar (in some degree) to anyone buying consumer goods. See Esther D. Reed, *The Limit of Responsibility: Engaging Dietrich Bonhoeffer in a Globalizing Era* (London: T&T Clark, Bloomsbury, 2018), esp. ch. 2.

45. Gen. 9:5-6 suggests that accountability before God for the taking of human life is absolute, i.e., not qualified or diminished in any way. In the global reset after the flood, God establishes universal standards, one of which is the prohibition against murder. The requirement is for reckoning (Heb. *adrsh*) for bloodshed at the hand of every human being (Heb. *umid eadm*). The verb *dāraś*, meaning to enquire or seek with care, is repeated to increase the sense of intensity of an investigative and insistent demand.

46. Not because of the relevance of the JWT to matters *mala in se* but because the JWT helps us understand why things are *mala in se* and why they cannot be done.

on LAWS or striving for regulation; future wars and armed conflict are inconceivable without this technology.)

The Stop Killer Robots campaign is clear that ‘we are crossing a moral line’, ‘no one would be safe’, and that ‘humans may fade out of the decision-making loop in certain military actions’.⁴⁷ Also clear is the ICRC concern that ‘their normal or expected circumstances of use, could not be sufficiently understood, predicted and explained’.⁴⁸ ‘The user of an autonomous weapon system does not choose the specific target, nor the precise time or place that force is applied. This process risks the loss of human control over the use of force and it is the source of the humanitarian, legal, and ethical concerns’.⁴⁹ On the other side, at the nation-state level, counter arguments have been voiced most clearly by the United States of America, notably in the U.S. Statement at CCW GGE Meeting: Intervention on Appropriate Levels of Human Judgment over the Use of Force, where the key issue identified was controllability of the weapon system and appropriate levels of human judgement over the use of force.⁵⁰ The United States focused on the human role in authorizing lethal force and controlling it: ‘States will not develop and field weapons that they cannot control’.⁵¹ Similarly, the understanding of ‘meaningful human control’ is problematized by those for whom machine autonomy using neural network training and ‘deep learning’ is understood to be under adequate human control. For McFarland and Galliot, for instance, human control is applied by militaries prior to firing.⁵² The problem of ‘meaningful human control’ is met, they imply, by the human control exercised in the training, pre-testing, exercise of the precautionary principle, defining of very restrictive and pre-approved attack targets etc., and the moral imperative is to engage technology in enacting accountability. ‘States will not develop and field weapons that they cannot control’.⁵³

Unpredictability has too often been practised as a political strategy for much hope to be placed in its utter undesirability for militaries.⁵⁴ Indeed, as Oliver O’Donovan has

47. Campaign to Stop Killer Robots, ‘Who Wants to Ban Fully Autonomous Weapons?’, 2021, <https://www.stopkillerrobots.org> (accessed 20 August 2021).

48. International Committee of the Red Cross, ‘ICRC Position and Background Paper: ICRC Position on Autonomous Weapons Systems’, Geneva, 12 May 2021, p. 7, <https://www.icrc.org/en/document/icrc-position-autonomous-weapon-systems> (accessed 20 August 2021).

49. International Committee of the Red Cross, ‘Autonomous Weapons’.

50. U.S. Mission to International Organizations in Geneva, ‘U.S. Statement at CCW GGE Meeting’.

51. U.S. Mission to International Organizations in Geneva, ‘U.S. Statement at CCW GGE Meeting’.

52. McFarland and Galliot, ‘Understanding AI and Autonomy’, pp. 51–52.

53. U.S. Mission to International Organisations in Geneva, ‘U.S. Statement at CCW GGE Meeting: Intervention on Appropriate Levels of Human Judgement over the Use of Force’, 16 November 2017, <https://geneva.usmission.gov/2017/11/16/u-s-statement-at-ccw-gge-meeting-intervention-on-appropriate-levels-of-human-judgment-over-the-use-of-force/> (accessed 24 August 2021).

54. James D. Boys, ‘The Unpredictability Factor: Nixon, Trump and the Application of the Madman Theory in US Grand Strategy’, *Cambridge Review of International Affairs* 34.3 (2021), pp. 430–51.

observed, a pressing concern over autonomous weapons is that these systems could be, or perhaps could later become, another example of the strategy of putatively removing responsibility from the actor.⁵⁵ Whether predicted and desired, or unpredicted and undesired, the categorical unacceptability of unpredictability is why there is *prima facie* moral attractiveness about the idea of a simple ban on making or possessing such weapons. Consider again the ‘fire, forget and find’ incident reported by the United Nations panel of experts on Libya with which we started:

On 27 March 2020, the Prime Minister, Faiez Serraj, announced the commencement of Operation PEACE STORM ... The enhanced operational intelligence capability ... allowed for the development of an asymmetrical war of attrition designed to degrade HAF ground unit capability ... The lethal autonomous weapons systems were programmed to attack targets without requiring data connectivity between the operator and the munition: in effect, a true ‘fire, forget and find’ capability.

Libya descended into political chaos and violence following the overthrow of President Muammar Gaddafi a decade ago. ‘The country became divided between two rival administrations, with the internationally recognized Government of National Accord (GNA) based in the West, while the self-styled Libyan National Army (LNA) controlled large areas in the East’.⁵⁶ The ‘fire, forget and find’ capability was exercised by forces backed by the government based in Tripoli against Libyan National Arab Army units under the command of Field Marshal Khalifa Haftar. The report from the panel of experts on Libya reports that a drone, described as ‘a lethal autonomous weapons system’ powered by artificial intelligence, hunted down and remotely engaged enemy militia fighters as they ran away from rocket attacks. Much remains unclear, notably, as Zachary Kallenborn observes, ‘whether the drone was allowed to select its target autonomously and whether the drone, while acting autonomously, harmed anyone’.⁵⁷ Whether or not the first time, it appears that the weapon operated on software-based algorithms ‘taught’ through large training datasets to classify objects, and that the weapon fired on targets ‘without data connectivity between the operator and the munition’ at the point of firing.⁵⁸ How predictably it selected targets is difficult to establish.

Consider further an incident reported by Jennifer Gibson, who works for Reprieve, a UK legal action non-governmental organization,⁵⁹ namely the case of Faisal bin ali Jaber whose brother-in-law and nephew were killed in a drone attack on the basis of suspicious

55. Private communication, 27 August 2021.

56. United Nations, ‘Secretary-General Underlines UN Commitment to Libya’, 23 June 2021, <https://news.un.org/en/story/2021/06/1094562> (accessed 20 August 2021).

57. Bulletin of the Atomic Scientists, ‘Zachary Kallenborn’, <https://thebulletin.org/biography/zachary-kallenborn/> (accessed 20 August 2021); Z. Kallenborn, ‘Was a Flying Killer Robot Used in Libya? Quite Possibly’, *Bulletin of the Atomic Scientists*, 20 May 2021, <https://thebulletin.org/2021/05/was-a-flying-killer-robot-used-in-libya-quite-possibly/> (accessed 20 August 2021).

58. Kallenborn, ‘Was a Flying Killer Robot Used in Libya?’.

59. For details see <https://reprieve.org/uk/about-us/>.

behaviour based on signals intelligence run through algorithms. Faisal's brother-in-law, Saleem, was known for speaking out forcefully in his sermons against Al Qaeda, and his nephew, Waleed, was a policeman. Shortly after one sermon, three men demanded to talk with Saleem. Suspicious that they might be Al Qaeda, Saleem took the meeting outside where he thought it would be safest. As Gibson says: 'Algorithms, at their best, merely tell us about relationships. They don't tell us whether Faisal's brother-in-law is meeting the men because he's planning an attack with them, or instead, if the meeting is to explain why Al Qaeda's ideology is wrong'.⁶⁰

Here, then, is the challenge: whether to lend support to political pressure only for an outright ban, or also to a legally binding international treaty for the regulation of weaponry that remains outside a ban? If the latter, questions about whether and/or how the JWT should evolve as *jus in silico* become pressing and two broad categories of question are in play. The first are quasi technical and relatedly ethical, for example: Was the problem due to the use of algorithms per se or the quality of the Social Network Analysis regarding the importance attached to connections rather than the content of those connections? Could these problems have been addressed ahead of time? Is the challenge now to adjust from thinking about weapons as singular instruments, for example a sword or pistol, to larger systems in which humans and machines interact? Hence the ICRC asks:

- What is the level of human supervision, including the ability to intervene and deactivate, that would be required during the operation of a weapon that can autonomously select and attack targets?
- What is the level of predictability and reliability that would be required, also taking into account the weapon's tasks and the environment of use?
- What other operational constraints would be required, notably on the weapon system's tasks, its targets, the environment in which it operates (e.g. populated or unpopulated area), the duration of its operation, and the scope of its movement?⁶¹

Such questions about degrees of autonomy and what might constitute the application of responsibility become critical if the JWT is to further investigate what an outright ban might mean, what weapons would be outlawed, and why, and what kind of regulatory regime might be viable. If increased adherence to the principles that underpin existing

60. Readers might recall the protest by Google employees that 'Google should not be in the business of war' – see S. Shane and D. Wakabayashi, 'The Business of War: Google Employees Protest Work for the Pentagon', *New York Times*, 4 April 2018, <https://www.nytimes.com/2018/04/04/technology/google-letter-ceo-pentagon-project.html> (accessed 20 August 2021). The company's Project Maven was customizing an AI surveillance engine using 'Wide Area Motion Imagery' for US Government drones to capture data that would detect vehicles and other objects, track their motions, and provide results to the Department of Defense. 'Letter to Google C.E.O.', *New York Times*, [no date], <https://static01.nyt.com/files/2018/technology/googleletter.pdf> (accessed 20 August 2021).

61. International Committee of the Red Cross, 'Autonomous Weapons'.

IHL (namely, humanity, necessity, distinction, proportionality) were potentially achievable using these weapon systems, their deployment might be deemed morally preferable to other systems. If, however, levels of unpredictability and unreliability were accepted or programmed-in, in effect, as a means of terrorizing the enemy, their use might be deemed abominable. Levels of predictability and reliability, and the degrees of autonomy and of detachment from selection of the target, are of utmost moral significance.

The related, second set of questions follows from the quasi technical and is overtly theologico-ethical: Has the JWT reached the end of the road because the detachment of war fighting from human agency is too great and robot-inflicted death intrinsically evil (*malum in se*), or can/should it be accepted that no fully autonomous weapons system exists because there are always human beings somewhere, and that evil is to be minimized (*minima de malis*) in the context of a justified war by ensuring that weapons systems are held to account? Should LAWS using neural network training and deep learning be regarded as inherently indiscriminate because using network training to select targets is unavoidably unpredictable,⁶² or under meaningful human control and subject to human agency by virtue of the programming, testing and subsequent accountability mechanisms? How is such a decision to be made?

‘Put your sword back into its place; for all who take the sword will perish by the sword’ (Matt. 26:52 NRSV). Jesus’ words were interpreted by Justin Martyr, Tertullian, and many others in the early Church, to preclude fighting against enemies and military service for Christians, and these words demand attention again. My (implicit) suggestion in this article is a rereading of these verses in this context of new weapons systems where a critical issue for weapons control is putting/not putting/not attempting to put decision-making and control beyond the reach of divine address. ‘The serpent deceived me, and I ate’ (Gen. 3:13). ‘The machine took its own decision and fired’. Is neural-network training or ‘deep learning’, in effect, an attempt to put decision-making beyond the reach of divine address? The divine command: ‘Put your sword back into its place’ is meaningful only when spoken to an agent capable of obeying. Are increasing levels of autonomy in weapons systems somehow an attempt to escape appropriate lines of accountability from agent to act and vice versa?

In addressing these questions, there is no denying that the distancing of human agency from the weapon at the point of lethality is the peculiar evil in question; loss of control obscures the line of accountability that leads back from the means to the agent. Nor is there any denying that autonomy is one of the emerging technologies identified by nation-states as a ‘key contributor to driving economic growth and delivering wide-ranging benefits for society’, and is being recognized by nation-states as having many potential applications within a defence context.⁶³ Automated drones are relatively cheap to purchase for

62. I.e., weapons that in their normal or expected circumstances of use cannot be directed at a specific military objective or whose effects cannot be limited as required by IHL. J.-M. Henckaerts and L. Doswald-Black, *Customary International Humanitarian Law Study Vol.1: Rules*, International Committee of the Red Cross (Cambridge: Cambridge University Press, 2005), Rule 71.

63. Gov.co.uk, ‘Guidance: Autonomy Programme’, 1 January 2018, <https://www.gov.uk/guidance/autonomy-programme> (accessed 20 August 2021).

potential adversaries and '[a]t root, autonomy is just a matter of programming the weapon to fire under given conditions, however simple or complex'.⁶⁴ The Christian pacifist and some JWT reasoners might choose to disengage because the distancing of human agency from lethality renders this weaponry categorically different from other weaponry. Others walking in the JWT (myself included) will locate the problem of agency squarely with human beings, within a human chain of accountability, and concentrate on giving that accountability meaningful legal force. From this position, to desire an outright ban and to prepare for a legally binding international treaty that sets out agreed parameters within which such weapons can be developed and used are not mutually exclusive. But those who recognize the need to prepare for the unavailability of an outright ban are called upon to develop the criteriological function of the tradition with respect to new weapons technologies because such will be required for any serious regulatory or, indeed, prohibitive attempt.

Future Work

All this matters, of course, because weapons control and control over weapons determines not only the way war is fought but also sometimes *whether* war is fought.⁶⁵ Armed conflict that is less and less visible to electorates and safer for military personnel is potentially more acceptable to governments. The 'liberal conscience' seems to require Western states to fight in ever more bloodless and 'humane' ways, with the knock-on erosion of *jus ad bellum* restraints and, indeed, for maintaining the ethical distinction between war and peace.⁶⁶ Until such time that de-weaponization is achieved, however, we must ask if the JWT is ready to develop new models of responsibility and accountability. Much of today's law of armed conflict and IHL has religious roots stretching back to the great twelfth-century Jewish rabbinic figure Maimonides, early Medieval church lawyers, comparable legal literature of the Islam, and beyond. Yet the recently renewed visibility of religion in public life, nationally and internationally, is yet to impact these debates. What, then, are the prospects for *jus in silico*?

Answers turn on whether new weapons technologies catastrophically undermine the JWT as a proposal for doing justice in the theatre of war. This in turn depends upon whether responsibility for actions performed with weapons, and the capacity to disarm them, is abrogated to the machine, and whether clear lines of accountability are in place for control over weapons. Working models of responsibility in deployment and accountability that run from procurement and design through political decision-makers

64. Z. Kallenborn, 'A Partial Ban on Autonomous Weapons Would Make Everyone Safer', *FP*, 14 October 2020, <https://foreignpolicy.com/2020/10/14/ai-drones-swarm-killer-robots-partial-ban-on-autonomous-weapons-would-make-everyone-safer/> (accessed 20 August 2021).

65. W. H. Boothby, *Weapons and the Law of Armed Conflict*, 2nd edn (Oxford: Oxford University Press, 2016), p. vii.

66. Christopher Coker, *Future War* (Cambridge: Polity, 2015), idem, *Humane Warfare* (London: Routledge: 2001); S. Moyn, *Humane: How the United States Abandoned Peace and Reinvented War* (New York: Farrar, Straus, and Giroux, 2021).

to shareholders and electorates is beyond our scope here. But it is important that we begin to see the issues and ask the necessary questions. Much work lies ahead. For the moment, we note that the Guiding Principles affirmed by the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapon Systems adopted by the 2019 Meeting of the High Contracting Parties to the Convention on Conventional Weapons affirm:

(b) Human responsibility for decisions on the use of weapons systems must be retained since accountability cannot be transferred to machines. This should be considered across the entire life cycle of the weapons system.

‘Human responsibility’ appears to be used expansively to mean that a person is morally and professionally answerable for actions performed and for knowing the purposes they seek to achieve. The implication appears to be that they are potentially subject to punishment for actions that are unlawful or otherwise contrary to the obligations attaching to their role. ‘Accountability’ pertains to the processes and norms that hold a person legally liable.⁶⁷ What this principle means in practice is the critical point at issue. In the ‘fire, forget and find’ incident in Libya mentioned above, who should be held to account for determining whether the act was unlawful? The Guiding Principles also state:

(a) International humanitarian law continues to apply fully to all weapons systems, including the potential development and use of lethal autonomous weapons systems.⁶⁸

As given, however, this principle does not test the proposition that existing IHL might not be sufficient to address new issues arising from the potential use of such weapons systems.⁶⁹ For instance, if LAWS using neural network training and deep learning are deemed to be at least potentially under meaningful human control and subject to human agency by virtue of the programming, testing and subsequent accountability mechanisms, should norms and standards be set for predictability and reliability in the achieving of discrimination and proportion? What peculiar challenges pertain to interpretation of the principles of humanity and military necessity under such conditions? Who should be held to account in the event that an act is found unlawful: the automated drone, the military commander, the politician who set the policy direction, the company or other body which sold the drone, the developer, the electorate (if appropriate)?

67. These directions in interpretation are confirmed at <https://documents.unoda.org/wp-content/uploads/2020/09/GGE20200901-Austria-Belgium-Brazil-Chile-Ireland-Germany-Luxembourg-Mexico-and-New-Zealand.pdf>, p. 5.

68. Group of Governmental Experts on Lethal Autonomous Weapons Systems (GGE LAWS), 2019, <https://documents.unoda.org/wp-content/uploads/2020/09/GGE20200901-Austria-Belgium-Brazil-Chile-Ireland-Germany-Luxembourg-Mexico-and-New-Zealand.pdf> (accessed 20 August 2021).


69. This point is made by the Group of Governmental Experts on Lethal Autonomous Weapons Systems.

Clearly, the actions for which each agent in this list might be individually accountable (i.e., attributable as its efficient cause or author) are different, and judicial processes would have to reflect these differences. Not all persons in the accountability chain will have fired the weapon. Not all will have charged its battery, set its course and loaded the ammunition. Not all will have signed off the requisite funding. Not all will have designed the control system on the weapon, its human interface, the algorithm which selects a particular target type or supplied the training data. Not all will have voted for the government that funded the initial research or invested in the company that prepared initial designs. All might somehow need to be included, however, in a full picture of what accountability for the (un)lawful taking of human life looks like in this situation—to which end, a discussion of imputation (L. *imputare* ‘enter in the account’, from *in-* ‘in, towards’ + *putare* ‘reckon’) is perhaps a next step for an integrated model of accountability wherein chains of accountability hold. In the meantime, the peril of imprecision with respect to the meaning and practice of accountability is that judicial processes can be sidelined and easily ignored, and the very possibility of doing justice in the theatre of war undermined.

Acknowledgements

I am grateful to Joseph Capizzi and Oliver O’Donovan for comments on an earlier draft—the former in the discussion of agency especially and the latter as cited in the text. Any errors remain my own. The position advanced is not necessarily theirs. Sincere thanks also to Anicée van Engeland for sustained conversation over many months.

ORCID iD

Esther D. Reed  <https://orcid.org/0000-0002-9714-6617>