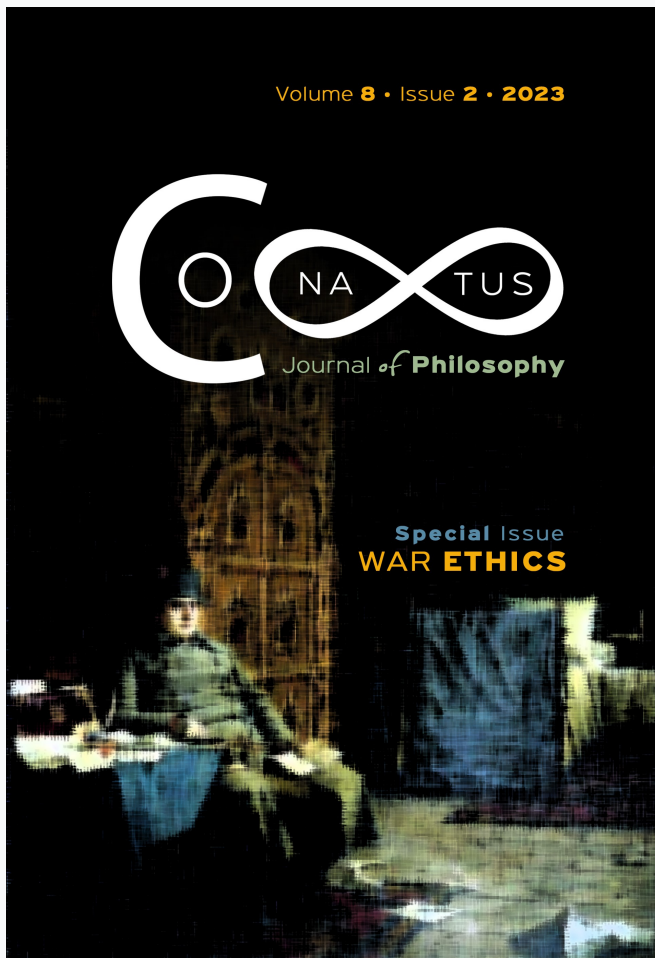


Conatus - Journal of Philosophy

Vol 8, No 2 (2023)

Conatus - Journal of Philosophy SI: War Ethics



An Ethic of Military Uses of Artificial Intelligence: Sustaining Virtue, Granting Autonomy, and Calibrating Risk

Nigel Biggar

doi: [10.12681/cjp.34666](https://doi.org/10.12681/cjp.34666)

Copyright © 2023, Nigel Biggar



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0](https://creativecommons.org/licenses/by-nc/4.0/).

To cite this article:

Biggar, N. (2023). An Ethic of Military Uses of Artificial Intelligence: Sustaining Virtue, Granting Autonomy, and Calibrating Risk. *Conatus - Journal of Philosophy*, 8(2), 67–76. <https://doi.org/10.12681/cjp.34666>

An Ethic of Military Uses of Artificial Intelligence: Sustaining Virtue, Granting Autonomy, and Calibrating Risk

Nigel Biggar

University of Oxford, United Kingdom

E-mail address: nigel.biggar@stx.ox.ac.uk

ORCID iD: <https://orcid.org/0000-0002-3005-4355>

Abstract

Artificial intelligence in military operations comes in two kinds. First, there is narrow or specific intelligence – the autonomous ability to identify an instance of a species of target, and to track its changes of position. Second, there is broad or general intelligence – the autonomous ability to choose a species of target, identify instances, track their movements, decide when to strike them, learn from errors, and improve initial choices. These two kinds of artificial intelligence raise ethical questions mainly because of two features: the physical distance they put between the human agents deploying them and their targets, and their ability to act independently of those agents. The main ethical questions these features raise are three. First, how to maintain the traditional martial virtues of fortitude and chivalry while operating lethal weapons at a safe distance? Second, how much autonomy to grant a machine? And third, what risks to take with the possibility of technical error? This paper considers each of these questions in turn.

Keywords: *artificial intelligence; war; weaponry; ethics; virtues; autonomy; risk; theological ethics*

I.

Artificial intelligence in military operations comes in two kinds. First, there is narrow or specific intelligence – the autonomous ability to identify an instance of a species of target, and to track its changes of position. Second, there is broad or general intelligence – the autonomous ability to choose a species of target, identify in-

stances, track their movements, decide when to strike them, learn from errors, and improve initial choices.

These two kinds of artificial intelligence raise ethical questions mainly because of two features: the physical distance they put between the human agents deploying them and their targets, and their ability to act independently of those agents. The main ethical questions these features raise are three. First, how to maintain the traditional martial virtues of fortitude and chivalry while operating lethal weapons at a safe distance? Second, how much autonomy to grant a machine? And third, what risks to take with the possibility of technical error?

II.

The feature of physical distance between military agent and military effect has given rise to worries about the future of traditional military virtues. Different cultures engender different kinds of military *ethos*, of course, and different *ethē* promote different sets of virtues. The military cultures of Jhinghis Khan's Mongols or the SS were not exactly the same as those of medieval Christendom or the British Army in the Second World War. Some virtues are bound by the nature of warfare to feature in all military cultures, most notably, physical courage, honour, and loyalty (I myself would add a certain kind of callousness.¹) To these generic military virtues, the specific military *ethos* of a Christianised culture will add charitable self-restraint and mercy. These Christian virtues are generated partly by a theological anthropology, according to which all humans share the status of sinners in need of divine forgiveness; and partly by a theological soteriology, according to which the punishment of wrongdoing should be in the service, never of the lust for vengeance, but only ever of a desire for "reconciliation" in the form of a just peace. These two theological doctrines issue in the following moral implications: that those who are morally justified in fighting should allow the ultimate end of a just peace to temper their military means; that those who wage unjustified war may not be regarded as simply morally alien; that the intention of just belligerency should not be to rid the world of evil by annihilating the unjust enemy, but rather to stop a particular outbreak of grave wrongdoing by rendering unjust warriors incapable of further fighting; and that there is no good reason to seek to harm non-combatants. These theologically generated moral implications entail that just warriors should cultivate the virtues of self-restraint and mercy in the manner of their use of lethal force.

¹ See Nigel Biggar, *In Defence of War* (Oxford: Oxford University Press, 2013), 117-119, 127, and 148.

Some ethicists believe that, by putting a human military operator of, say, a semi-autonomous, armed unmanned aerial vehicle (UAV) or “drone” at a safe arm’s length from the battlefield, artificial intelligence tends to corrode military virtues.² I am not persuaded. It is true that a uniformed agent in Arizona or East Anglia, who is operating a drone in Afghanistan, is completely safe from physical harm, and therefore does not have to exercise the courage necessary to overcome the natural fear of such harm. However, that is only because of the happenstance that the enemy in Afghanistan lacks the ability to strike back with long-range missiles. Operating a military drone over Russia would not be quite so safe. Besides, the virtue of physical courage is a typical requisite of front-line combat troops – and of support troops who might find themselves pushed into the front line. It is not typically requisite of those who, though civilian, are nevertheless contributing to the waging of war safely remote from the front line. That is to say, the waging of war involves a spectrum of exposure to physical harm – as it has probably always done – whereby some war-wagers are safer than others. That is to say, the virtue of physical courage has not been expected of *all* war-wagers – let us call them “warriors” – for a long time, perhaps ever.

Robert Sparrow observes that, while the pilots of UAVs lack the opportunity to exercise and cultivate physical courage, they can still exercise and develop moral courage, whether in deciding to take human life or in refusing to obey what appears to be an illegal or immoral order. And the serious cost of bearing the responsibility for exercising such courage is evident in reports among Predator and Reaper pilots of PTSD. But he worries that this does not distinguish them from ambulance drivers, surgeons, and rescue workers, except insofar as their role involves a deliberate decision to kill. And in that respect, it does not distinguish them at all from armed policemen.³ To which my response is: but why should it?

As for the virtue of honour in the general sense of upholding the standards of conduct expected of members of the military profession or unit, Sparrow rightly observes that UAV operators are less likely to be thrown off the moral course by fear of death or injury than combatants.⁴ Sometimes, however, military honour is perceived specifically in terms

² For example, Robert Sparrow, “War without Virtue?” in *Killing by Remote Control: The Ethics of an Unmanned Military*, ed. Bradley Jay Strawser, 84-105 (Oxford: Oxford University Press, 2013), 88.

³ *Ibid.*, 89, and 94.

⁴ *Ibid.*, 97.

of chivalry, and chivalry in terms of fairness. Accordingly, it seems dishonourable that a UAV operator should be able to strike the enemy with devastating force, while remaining absolutely immune from retaliation. The gross asymmetry of power seems grotesquely unfair. This is a common perception.⁵ But it is a mistaken one. The aim of any belligerency is so to overwhelm the enemy as to disable him from continuing to fight. This is done by applying the greatest possible force against him at his weakest point. Whatever the place of fairness in war, it does not consist of making sure that the enemy is equally well resourced before one engages him. Sparrow is largely correct, therefore, when he writes, “we need to be careful to avoid relying on an argument about chivalry here. War is not a game, and there is no reason that it should be fair.”⁶

Regarding loyalty, there are different kinds and not all kinds should be expected of all warriors. For a Christian, of course, there can be no such thing as absolute loyalty to any human institution, since the Christian’s primary loyalty must be to God and his moral law, and since human institutions sometimes transgress that law. As Sir Thomas More said on the scaffold moments before he was beheaded, “I die the King’s good servant, but God’s first.”⁷ Members of a combat unit need to be able to depend on their comrades to protect them and aid them in the most threatening and terrifying of circumstances, if they are to be militarily effective. That will require group loyalty of a peculiar intensity. Other warriors will need to show themselves loyal – under God – to a just cause, loyal to the state that fights in a just cause, and loyal to the state institutions and military units that serve that just cause. But they need not cultivate the same kind of loyalty as a combat unit. Again, Sparrow worries that this blurs the line between civilian and military.⁸ But I fail to see why that is a problem.

Concerning the virtue of charitable self-restraint, it seems obvious that military agents who are distanced from the confusing, threatening

⁵ In my book, *Colonialism: A Moral Reckoning* (London: HarperCollins, 2023), I report two cases where historians (Dan Hicks, and William Beinart) think that the use of Maxim guns and naval artillery by the British against native Africans in what is now Rhodesia and Benin was morally objectionable, because the balance of military power was so unequal.

⁶ Sparrow, 99.

⁷ According to a contemporary report carried in the *Paris Newsletter*. See Nicholas Harpsfield, *The Life and Death of Sir Thomas Moore, Knight, Sometyms Lord High Chancellor of England*, ed. Elsie Vaughan Hitchcock and Raymond Wilson Chambers (London: Oxford University Press, 1932), Appendix III, 266: “Après les exhorta, et supplia tres instamment qu’ils priassent Dieu pour le Roy, affin qu’il luy vouldist donner bon conseil, protestant qu’il mouroit son bon serviteur et de Dieu premierement.”

⁸ Sparrow, 97.

maelstrom of the battlefield, and whose security against risk permits maximal caution, are more likely to be capable of exercising restraint than combat troops. What is more, according to Dave Grossman's 1995 book, *On Killing: The Psychological Cost of Learning to Kill in War and Society*, the closer troops are to the enemy, the greater their reluctance to kill.⁹ To this Paul Scharre adds the observation that the cameras of a UAV can bring its pilot face-to-face with the target.¹⁰ Therefore, while that may be responsible for causing the unhappy effect of PTSD, when he decides to kill, it is also likely to cause the happy effect of increasing his reluctance to make such a decision. For this reason, I doubt Sparrow's argument that, because UAV pilots never *meet* their enemies, such compassion as they have "must necessarily be abstract, which will also rule out genuine acts of mercy."¹¹ Meeting the enemy may not be necessary to induce merciful restraint in killing; seeing them may suffice.

In general, I am sceptical that the military uses of artificial intelligence will lead to a decline in military virtues.¹² As the means of war evolves, so do the relevant virtues and their distribution. While the traditional virtues will still be required of military personnel performing traditional roles, there may be novel roles that require a different set of virtues. What will be important, however, is not to require a person who has been made to cultivate one set of virtues to perform a role that requires a different set.

I do not agree, therefore, with Shannon Vallor, when she argues that the military use of artificial intelligence will generally deskill military personnel, depriving them of the opportunity to cultivate through experience the virtue of practical wisdom (or prudence), which is needed for making the right choices in rapidly changing circumstances about "who or what gets targeted, or when, in which circumstances, or with what degree of force."¹³ For sure, the pilots of UAVs will not develop the vir-

⁹ Dave Grossman, *On Killing: The Psychological Cost of Learning to Kill in War and Society* (Boston: Little, Brown, and Co., 1996), Section III, "Killing and Distance: From a Distance, You Don't Look Anything Like a Friend."

¹⁰ Paul Scharre, *Army of None: Autonomous Weapons and the Future of War* (New York: W. W. Norton & Co., 2018), 275-276.

¹¹ Sparrow, 102.

¹² I observe that Sparrow agrees: "it is doubtful that wars will ever be fought entirely by weaponry that eliminates the need for the traditional martial virtues." *Ibid.*, 105.

¹³ Shannon Vallor, "The Future of Military Virtue: Autonomous Systems and the Moral Deskilling of the Military," in *2013 5th International Conference on Cyber Conflict: Proceedings*, eds. Karlis Podens, Jan Stinissen, and Markus Maybaum, 471-486 (Tallinn: NATO Cooperative Cyber Defence Centre of Excellence Publications, 2013), 478, and 480.

tue of physical courage as must those of manned aircraft, together with combat soldiers and sailors. However, being safely removed from the theatre of operations, UAV pilots are less likely to have their practical judgement thrown off course by pain or fear or anger, and over time they will accumulate experience in decision-making and thereby cultivate prudence. In other words, they will be stronger in one military virtue, while being weaker in another – differently skilled, not de-skilled.

III.

The second ethical question raised by the military uses of artificial intelligence is how much autonomy to grant weapons, and this in turn raises a further issue about virtue. The pressure to increase autonomy arises partly because of the danger that the communication-link with a weapon might be broken and partly because of the need for speed in responding to enemy action. Speed, and therefore autonomy, is especially important for effective cyber-defence.¹⁴

Autonomy comes in degrees, and is never absolute. According to Sparrow, almost all of the “robotic” weapon systems currently being developed are either remotely operated or unmanned, rather than fully autonomous.¹⁵ Their autonomy consists of using sensors to read the environment and identify a target, and then processors to decide how to respond, say, by adapting to the target’s movements.¹⁶ Beyond that, however, a human operator is usually required to make key decisions or at least has the power to intervene in the machine’s decision-making process. That is to say, humans remain either “*in the loop*” or “*on the loop*.” The key decision that carries the greatest moral weight is, of course, the decision to strike, and according to Paul Scharre, “[f]or most weapons systems in use today, a human makes the decision whether or not to engage the target.”¹⁷ But not all systems. Israel’s Harpy drone, for example, is largely autonomous: it not only loiters overhead and searches for potential targets but can decide to strike without asking for human permission.¹⁸ Yet even here, the drone’s autonomy is not absolute: the human operator still determines which *species* of target the drone should home in on – say, enemy radars – while the drone itself decides only which *specimens* to attack.¹⁹

¹⁴ Scharre, 216.

¹⁵ Sparrow, 86.

¹⁶ Scharre, 41-42.

¹⁷ *Ibid.*, 44.

¹⁸ *Ibid.*, 5, 46, 48, and 64.

¹⁹ *Ibid.*, 48.

It is autonomy over the decision to strike a target that raises moral issues. For that decision, which may cause the grave non-moral evil of the destruction of human life, ought to issue from deliberation about the “just war” principles of discrimination and proportionality. Applying these principles on the battlefield is not a straightforward, logical, mechanical operation. It requires the interpretation of circumstances, the estimation of military necessity and the urgency of action, and perhaps the discrimination of combatants in civilian clothes from non-combatants. However, whereas computers are often more intelligent and faster in performing narrowly specified tasks, according to Scharre, they “still fall far short of humans in understanding context and interpreting meaning,” that is, in “general intelligence.”²⁰ And artificial general intelligence is currently only “a hypothetical future.”²¹ “Unlike humans, autonomous weapons systems lack the ability to step outside their instructions and employ ‘common sense’ to adapt to the situation at hand.”²² Whereas human agents are “capable of using their common sense and better judgment to comply with the intent behind a rule, rather than the rule itself,”²³ autonomous systems are not. What this implies is that we cannot expect a weapons system to exercise the virtue of prudence, and that we should expect a fully autonomous system, which cannot be recalled or supervised and which can make a decision to strike on its own, to act imprudently.

Scharre suggests that an autonomous weapon could observe the principle of proportionality, if humans programmed it to avoid risking the lives of a certain number of non-combatants.²⁴ But that would be to employ a very crude utilitarian understanding of the principle. According to classic “just war” thinking, provided that one does not intend to harm non-combatants, and provided that one actualises that intention by earnestly seeking to avoid causing such harm, how much

²⁰ Ibid., 6, and 95.

²¹ Ibid., 231.

²² Ibid., 146.

²³ Ibid., 308. A famous example of this is when Commodore Horatio Nelson disobeyed the orders of Admiral Sir John Jervis at the Battle of Cape St. Vincent in 1797. Nelson’s biographer, John Sugden, comments that Nelson “prided himself on what he called ‘political courage,’ and repeatedly acted on it, even in contravention of the orders of superiors.” Yet, “if Nelson acted against the strict letter of Jervis’ orders he most assuredly remained within their spirit.” See John Sugden, *Nelson: A Dream of Glory* (London: Pimlico, 2012), 695. It seems that the admiral agreed. For, when his flag captain complained to him about Nelson’s disobedience, Jervis is said to have responded, “It was certainly so, and if you ever commit such a breach of orders, I will forgive you also.” (Ibid., 706).

²⁴ Scharre, 257.

risk one may take with non-combatant lives will depend on a range of circumstances. These will include the importance of the military objective, the military possibility and affordability of adopting less risky ways of achieving it, and the political consequences of non-combatant deaths. The principle of proportionality requires that risks to life be calibrated to a set of circumstances, and since circumstances are constantly changing and not all sets of them can be predicted, there does not exist an absolute number that can be programmed into a weapons system that would make its action proportionate.

Scharre also suggests that it would be morally safe to use autonomous weapons systems in an environment devoid of civilians.²⁵ That would certainly avoid imprudence causing a disproportionate number of non-combatant deaths. But the principle of proportionality also applies to the killing of enemy combatants: one should not kill more of them than military necessity requires. And Scharre himself makes the point that autonomous weapons would find it difficult to recognise genuine attempts at surrender, since that requires discerning intent amidst circumstances that might be highly ambiguous.²⁶

IV.

One might conclude that, morally speaking, one should never permit a weapons system to be fully autonomous in the sense that it can make the decision to strike on its own and without suffering interference from a human supervisor. The risk of disproportionate deaths, both combatant and non-combatant, would be too high. Yet, risks of some kind or another are often unavoidable, and their proportionality varies according to circumstances. The graver the threat, the higher the risks worth taking. So, there may be grave circumstances, where launching fully autonomous weapons is proportionate.

However, for such risky action to be prudent, those deciding upon it would have to have their eyes fully open. The temptation, especially with novel, sophisticated technology, is to indulge in wishful thinking and to downplay the risks.²⁷ In addition, there is the phenomenon of “automation

²⁵ *Ibid.*, 257.

²⁶ *Ibid.*, 259-260.

²⁷ The roboticist, Ron Arkin, expresses such over-confidence in technology. See Sharon Vallor, “The Future of Military Virtue,” 480; Scharre, 280, and 282-283; Brian Stiltner agrees: “Hyperbolic rhetoric surrounds new weapons. Political and military leaders often excitedly claimed that a new weapon is going to make a decisive difference or end a war. Almost always they overpromise.” See Brian Stiltner “A Taste of Armageddon: When Warring is Done by Drones and Robots,” in *Can War Be Just in the 21st Century? Ethicists Engage the Tradition*, eds. Tobias Winwright and Laurie Johnston, 14-28 (Maryknoll, NY: Orbis, 2015), 20.

bias,” that is, human deference to machines.²⁸ Yet, as Scharre rightly says, “100 percent error-free operation is impossible” and “[f]ailures are inevitable in complex, tightly coupled systems, and the sheer complexity of the system inhibits predicting when and how failures are likely to occur.”²⁹ Therefore, before launching fully autonomous weapons, the morally responsible human agents need to stare the worst-case scenarios squarely in the face and satisfy themselves that they are worth risking, and that, should they come about, they could be afforded.

In some cases, the cost will not be affordable and so the risk not worth taking. If the price is military defeat, then that should be borne. The tradition of “just war” thinking sanctions belligerency only under certain conditions. Absent those conditions, war is not just. At that point, the “just warrior” clanders off his war-horse and joins the pacifist on his knees, praying God to secure the justice that he cannot. Then, together, they rise and look around for non-military means of resistance.

References

- Biggar, Nigel. *Colonialism: A Moral Reckoning*. London: HarperCollins, 2023.
- Biggar, Nigel. *In Defence of War*. Oxford: Oxford University Press, 2013.
- Grossman, Dave. *On Killing: The Psychological Cost of Learning to Kill in War and Society*. Boston: Little, Brown, and Co., 1996.
- Harfsfield, Nicholas. *The Life and Death of Sir Thomas Moore, Knight, Sometyes Lord High Chancellor of England*, edited by Elsie Vaughan Hitchcock and Raymond Wilson Chambers. London: Oxford University Press, 1932.
- Human Rights Watch. *Losing Humanity: The Case against Killer Robots*. New York: Human Rights Watch, 2012.
- Scharre, Paul. *Army of None: Autonomous Weapons and the Future of War*. New York: W. W. Norton, 2018.
- Sparrow, Robert. “War without Virtue?” In *Killing by Remote Control: The Ethics of an Unmanned Military*, edited by Bradley Jay Strawser, 84-105. Oxford: Oxford University Press, 2013.
- Stiltner, Brian. “A Taste of Armageddon: When Warring is Done by Drones and Robots.” In *Can War Be Just in the 21st Century? Ethicists Engage the Tradition*, edited by Tobias Winwright and Laurie Johnston, 14-28. Maryknoll, NY: Orbis, 2015.

²⁸ Scharre, 144-145, 170, 278-279, and 324-325.

²⁹ *Ibid.*, 151, and 154.

Sugden, John. *Nelson: A Dream of Glory*. London: Pimlico, 2012.

Vallor, Sharon. "The Future of Military Virtue: Autonomous Systems and the Moral Deskillling of the Military." In *5th International Conference on Cyber Conflict: Proceedings*, edited by Karlis Podens, Jan Stinissen, and Markus Maybaum, 471-486. Tallinn: NATO Cooperative Cyber Defence Centre of Excellence Publications, 2013.