

THE MORALITY OF AUTONOMOUS ROBOTS

Aaron M. Johnson* and **Sidney Axinn****

**Electrical and Engineering Department, University of Pennsylvania, Philadelphia, United States;*

***Philosophy Department, University of South Florida, Tampa, United States*

While there are many issues to be raised in using lethal autonomous robotic weapons (beyond those of remotely operated drones), we argue that the most important question is: should the decision to take a human life be relinquished to a machine? This question is often overlooked in favor of technical questions of sensor capability, operational questions of chain of command, or legal questions of sovereign borders. We further argue that the answer must be 'no' and offer several reasons for banning autonomous robots. (1) Such a robot treats a human as an object, instead of as a person with inherent dignity. (2) A machine can only mimic moral actions, it cannot be moral. (3) A machine run by a program has no human emotions, no feelings about the seriousness of killing a human. (4) Using such a robot would be a violation of military honor. We therefore conclude that the use of an autonomous robot in lethal operations should be banned.

KEY WORDS: Autonomous robots, morality, human dignity, military honor

Introduction

As technology evolves and robotic systems become more and more autonomous, the prospect of an autonomous robotic war-fighter becomes a real possibility. Unmanned, remotely operated air and ground robots have been in routine use throughout the current wars in Afghanistan and Iraq. It is no secret that while official policy states that these robots will retain a human in the control loop, at least for lethality decisions, this policy will change as soon as a system is demonstrated that is convincingly reliable (Sharkey 2010). The number of nations developing this technology is growing – a recent report holds that 'close to 50 nations now possess the same unmanned technology [that the US uses]' (Singer 2012: SR5), and the appearance of autonomous robots may not be far behind.

Before we get to that point, society needs to consider the moral implications and extend the laws of war to more explicitly deal with these new kinds of weapons. The Geneva Conventions¹ of 1949 were signed just three years after ENIAC (the first general purpose computer) debuted and their framers could not possibly have imagined the rapid advancements in integrated circuits or artificial intelligence of the past 60 years. Similarly, the moral implications of nuclear weapons were not publicly debated until after their first use, and many of the scientists who worked on the Manhattan Project later regretted ignoring those moral issues (Szilard 1945). Therefore we must explore the moral issues related to autonomous robots now before they are fielded.

The first of Isaac Asimov's (1950) well-known 'Three Laws of Robotics' was that a robot may not injure a human being, or through inaction, allow a human being to come to

harm. When that assertion first appeared, many of us understood and agreed with it. That our own machines may harm us seemed quite wrong. Now we have the possibility of robots *designed* to cause injury and death to humans. War is a special (but not unusual) situation, and the matter of restrictions on robots must be carefully considered.

For this study, we will separate two classes of unmanned (aerial, ground, underwater, etc.) systems. What are called 'unmanned drones' have no human on board, but have a human in control, in the loop. These include remotely piloted aerial vehicles such as the Predator and weaponized ground robots such as the Talon SWORDS. Unmanned systems may have some autonomy, in navigation or target tracking; however, a human is still remotely watching a computer screen and making the final decision on when and what to fire upon. The unmanned drone simply allows a human to induce action far away from his/her physical location, either around the corner or around the globe.

'Autonomous robots' refers to systems that are directed by a computer program alone, with no human in the loop. Autonomous robots may make targeting decisions and fire weapons against humans or other targets on their own based on some computer algorithm. They may be given a task by a human commander, such as patrol this area or find a fugitive; however, they process their sensor feedback without human assistance and decide if and when they should fire upon a target. The focus of this paper is on this second category of systems, where a human is no longer making the lethality decision. Non-lethal autonomous robots, that is robots that either do not possess deadly firepower or those that do not use it autonomously, may bring up other issues that are outside the scope of this paper.

There are many potential issues with using autonomous robotic weapons, such as whether their algorithms can be discriminatory enough to avoid killing civilians. These are significant technical questions, but do not approach the importance of the key moral issue: should they be allowed at all? Regardless of their sensor technology and programmed algorithms, should we send robots on to the battlefield to kill our human enemies? Until that question has been answered, the details of a given robotic system or its operational scenarios is secondary.

Laws of War

Before we explore the issue of autonomous robotic weapons systems, we will first summarize the existing laws of war.² There are written and unwritten laws of warfare, and almost every country in the world has agreed to them. In written form, they consist of the Geneva Conventions and their amendments, and the earlier Hague Rules.³ In unwritten form, the US Army⁴ (Department of the Army 1956: 4) explains 'this body of unwritten or customary law is firmly established by the custom of nations . . .'. The manual adds that the customary laws of war may be compared with 'the unwritten Anglo-American common law.' The force of treaties like the Geneva Conventions should be clear and well known. For example, under the US Constitution, Article VI, clause 2, such treaties are part of the supreme Law of the Land. The army's (Department of the Army 1956: 7) recognition of this specifies that such treaties 'have a force equal to that of laws enacted by the Congress.'

Why are laws of warfare needed? On the first page of FM 27-10 (Department of the Army 1956), the purposes are listed as: (1) 'protecting both combatants and noncombatants from unnecessary suffering' (explained as suffering not needed for the military objective); (2) 'safeguarding certain fundamental human rights of persons who fall into the

hands of the enemy . . .'; and (3) 'facilitating the restoration of peace.' Without these laws, war would be more brutal than needed, and restoring peace would be more difficult.

Like all laws, these are broken on occasion. Laws may not always stop the forbidden acts from happening; rather they tell us at what one ought to be shocked. They are a common agreement of what should not occur, and therefore at what to be surprised or shocked. This 'shock' can range from raising an eyebrow to violent outrage, while the laws can range from 'no parking in front of a fire hydrant' to a ban on murder. That the Geneva and Hague Conventions are not followed meticulously is not an argument for ignoring them. Military honor requires fighting under the requirements of these conventions, even when your enemy does not.

Kinds of Weapons

To better understand the implications of unmanned and autonomous weapons we will consider a sampling of existing weapons and the issues and laws related to them. There is a large spectrum of conventional weapons, with varying moral and legal standings:

1. Some objects that are not designed as weapons, such as toasters, are still capable of inflicting harm (by electric shock, for example). Since their design and intended use are not lethal, any design defect that causes harm could fall under criminal negligence laws, and any misuse would not be blamed on the toaster. Similarly, for all weapons systems, including unmanned drones or autonomous weapons, design or programming bugs can be grounds for prosecuting the manufacturer, while improper deployment or use is clearly the fault of the operator. The precedent of both requiring rigorous testing, but also not exempting manufacturers of liability, does not change with an autonomous weapon, and conventional laws and customs still apply.
2. There are a number of conventional weapons, including guns, dumb bombs, spears, and so on, that are widely accepted as legal provided that they are used against a legitimate target under the laws of war. This requirement would obviously still apply to any unmanned or autonomous system.
3. Some weapons such as guided missiles (heat seeking, Global Positioning System (GPS) guided, Patriot) are semi-autonomous, in that they can adjust their trajectory in flight. We consider these to be in the same category as dumb bombs as they are still launched and targeted by humans; they just have better aim. Similarly, unmanned drones are extensions of conventional weapons – the human operators are still launching and targeting the weapons systems; they are just controlled over a greater distance.
4. The use of defensive weapons, such as electric fences, is well understood under the laws of war, assuming that they are used to secure a perimeter and are well marked. Any perimeter breach can be considered an act of hostility.
5. There are a number of contentious weapons, such as landmines. CCW⁵ lists a lot of technical requirements (mark, record, and clear) for anti-personnel mines that get around the discrimination issue and longer-term effects. In addition, most nations have gone further and banned them outright in the Ottawa Convention,⁶ on the grounds that they are designed to cause pain, and that the technical requirements are often ignored. Note, however, that the United States and several other major military powers are not party to the Ottawa Convention.

6. Chemical/gas weapons have a unique history. They were once only allowed in retaliation in the US position (Department of the Army 1956: Change No. 1). Since the United States signed the Chemical Weapons Convention of September 1992, chemical weapons are now not allowed in either a first-strike or second-strike situation.
7. A large number of other weapons are banned outright, including barbed lances, blinding lasers, and biological weapons. These are completely banned due to the excessive suffering they cause. Even if, for example, biological weapons are used against a nation, current agreements still bar that nation from retaliating with them.
8. Nuclear weapons maintain a strange status where they are generally considered unethical due to excessive suffering, civilian casualties, and lasting effects, although they are not specifically banned under any treaties or conventions. At this point, in the Advisory Proceedings of the International Court of Justice, speaking for the United States, John McNeill of the Department of Defense stated that 'the law of armed conflict governs the use of nuclear weapons – just as it governs the use of conventional weapons.' The court went on to rule that the threat or use of nuclear weapons would violate existing laws (Burroughs 1999). Therefore, since such weapons cannot spare civilians and they cause unnecessary suffering, there is a *de facto* ban, apparently.

The current legal status of autonomous weapons is almost certainly the same as nuclear weapons once held, that is not banned, 'in the absence of any customary rule of international law' (Department of the Army 1956: 18), though we argue that the laws of war should be amended to account for this new technology.

Autonomous Weapons

Even though there is not currently any international law explicitly governing unmanned or autonomous weapons, this does not mean that their use is completely unrestricted. If a soldier is in control when a war crime is committed, that soldier and his superiors may be court-martialed. The chain of command is well established for all weapons systems, including remotely piloted drones. However, when the robot is controlled by a program, not a human, the usual court-martial pattern may not be usable. One of the questions that has been raised is: where is moral responsibility to be placed for a robot controlled by its own program alone (Sparrow 2007, Human Rights Watch 2012)? This is really a detail, and can be answered using the same rules that apply to any weapon system. If the robot should not have been launched, the chain of command can still be well established, including who launched the drone, who ordered the launch of the drone, and so on. If it is a matter of a robot mistake, for instance due to a manufacturing or programming error, then the same rules that govern quality assurance and verification of toasters or conventional weapons should apply. If the weapons misfired due to a design defect, the manufacturer can be held liable. Furthermore, these weapons are able to record their sensory input and command decisions, meaning that any situation could be fully reviewed after the fact. These questions focus on responsibility after a war crime has been committed, but ignore the question of whether the system should have been deployed in the first place.

Much attention has also been given to the question of discrimination, that is whether a program can distinguish between a legitimate target, a belligerent, and an innocent civilian, as well as choosing the magnitude of force required for the military objective

(Sharkey 2008, Singer 2009, Human Rights Watch 2012). Surely any system considered for deployment would first be subjected to thorough evaluation. We assume that the work on this will eventually be successful and that programs will become sophisticated enough to pass any test of discrimination ability (if such a test could truly be made) at least as well as humans do (Arkin 2009). That does not necessarily mean that they will be more discriminate than humans in all situations, but rather any specific test of discrimination will certainly be passed eventually as technology continues to evolve. In this paper we are concerned with the issues inherent to using autonomous weapons, and not the limitations of current technology.

Another issue that has been brought up is the induced risk to human soldiers due to a drone or robot malfunction, for instance if it crashes behind enemy lines and a rescue team must be sent in after it, or if their efficacy and low risk causes nations to more easily go to war (Singer 2009, Ministry of Defence 2011, Human Rights Watch 2012). Furthermore, the lack of a human pilot in both drones and robots reduces the risk in using these weapon systems without a declaration of war, and according to US Rep. Dennis J. Kucinich: 'American drones in Pakistan have killed as many as 3,378 people. . . . We've not declared war on [Pakistan], but our weapons have killed innocent civilians' (Kucinich 2012: H6377). These are all crucial tactical and strategic issues of when is it legal or beneficial to use such weapons, but do not affect the ethical question of whether they should be used at all.

Some worry that using either unmanned or autonomous technology means that an army would 'have no skin in the game,' and implicitly that a war is waged justly (among other matters) only if there exists potential for damage to either side. However, just war theory does not require equal power or risk on each side.⁷ Increased range is a natural progression of technology and there is nothing inherently immoral in the fact that these weapons can strike from a distance. Similar arguments have been made for centuries about other distance weapons such as the crossbow, which was banned by Pope Innocent II. Is it wrong to use a crossbow to attack an enemy before they can get close enough to strike you with their sword? No, and similarly the distance between an unmanned drone operator and target does not in itself bring up any moral issues. Even if there is no direct threat to the remote operators or commanding officers at the time of attack, they certainly still have psychological skin in the game, they have chosen the target, and the army or nation as a whole is still vulnerable. We will revisit the question of personal risk later when we consider the concept of honor.

Proponents of autonomous technology cite their potential tactical benefits and argue that they reduce the risk of human harm (at least on one side of the battle), and we are therefore morally compelled to use them if available and efficacious (Strawser 2010). There is little doubt that an autonomous robot could help an army, just as in the past nations have sometimes used other contentious weapons to their advantage. Any time a new weapon is developed that allows you to more easily kill or hurt your enemy without being killed yourself there is a tactical benefit to using the weapon – for example, the US deployment of atomic weapons at the end of the Second World War had tactical benefits that are said to have led to the Japanese surrender. But as clearly established in the precedents of atomic/nuclear weapons, chemical weapons, anti-personnel land mines, barbed spears, and so on, the efficacy of a weapon is not justification for its use.

The Major Question

All of these questions ignore the bigger, more important moral question: should we relinquish the decision to kill a human to a non-human machine? In this paper we argue primarily that the discussion needs to focus on this most important question, and not on the details of the technology or its efficacy. We now further argue that the decision to take a human life must be an inherently human decision and that it would be unethical to allow a machine to make such a critical choice. The concept of what a rational human is, a being that can give itself the moral law, is essential to considering this matter.

The Concept of 'Human'

The difference between a human and a machine calls for some analysis. The history of thought has dealt with this in several ways, and among others Immanuel Kant comes to mind. Famously, for him humans have dignity; mere objects have a price. Humans are to be treated as ends in themselves; mere objects may be used as means to our ends (Kant 1959: 429). Christopher Coker (2008: 151) warns that 'Robots will inhabit a world of means largely divorced from ends.'

A mouse can be caught in a mousetrap, but a human must be treated with more dignity. A mousetrap kills targets with certain characteristics based on certain behavior, that is anything of sufficient mass eating or at least touching the bait. The trigger is designed to attack based on the mousetrap's perception of the target and its actions. The complexity of the trigger is not what we are concerned with – a mouse can be killed by a machine, as it has just a very small amount of dignity. A robot is in a way like a high-tech mousetrap; it is not a soldier with concerns about human dignity or military honor. Therefore, a human should not be killed by a machine as it would be a violation of our inherent dignity.

The concept of human dignity is complicated and explained variously. As a minimum, although not without controversy, we may take the statement of the Universal Declaration of Human Rights (1949, article 1): 'All human beings are born free and equal in dignity and rights.' Dignity is often taken as the property that makes humans eligible for the human rights listed in the Universal Declaration. Put another way, in Kant's (1959: 435) phrase, dignity means that the individual has 'an intrinsic worth,' and has 'no equivalent.' This is to say that each human must be respected for his or her unique inherent or intrinsic value.

It is widely accepted that all humans have a certain kind of equality; they are judged morally by the same rule. The American Declaration of Independence puts the attitude toward all humans in a vivid way. Near the end we find the statement, we hold our British brethren 'as we hold the rest of mankind, Enemies in War, in Peace Friends.' All humans are potentially our friends, and they all deserve to have us respect their inherent dignity.

Is there a loss of dignity when a human fights with a machine, compared to fighting with another human? As a non-lethal game it is acceptable;⁸ but in a fight to the death the matter is different and far from trivial. To give a programmed machine the ability to 'decide' to kill a human is to abandon the concept of human dignity. Humans are sometimes accidentally killed by machines, but for an autonomous robot/drone to be programmed to kill a human is to treat a rational being as if it were merely an object.

Morality

Humans give themselves moral commands; programs are given the commands they are to follow. A machine can replicate an action that has been called moral, but morality does not come from following someone else. In this way a machine can *act* morally, by mimicking its programmer, but it cannot *be* moral. Similarly, when playing back a video of a moral act, one would not say that the video was moral; it is simply replaying the moral act of its subject. An autonomous robot is clearly more complicated than a video, and can more directly interact with the world and respond to feedback; however, it is still just playing back moral actions, and is not truly making the decisions for itself.

Moral commands are based on values, and values are produced and indicated by sacrifices (Axinn 2010, also see section on military honor, below). While robots are aware of following orders, they are not aware of making sacrifices. Artificial intelligence still has no real notion of sacrifice (and artificial morality is still just a phrase and not a developed subject). Therefore robots have no values of their own, although they are following the values of their programmers. A distinctive human characteristic, therefore, is the ability to think morally based on one's values, and to give oneself the moral commands.

Morality, according to Kant, requires that the principle of one's action must be one that is capable of being followed universally. Again, not the specific act, but the *principle* of the act must be one that everyone can follow. The ability to judge and extract the principles of an action cannot be codified or programmed. As Kant (1965: A133) put it, 'judgment . . . cannot be taught.' Furthermore, Kant (1965: A135) insists, 'general logic can supply no rules for judgment. . . .' Without this ability, we find once again that a robot can only mimic morality by replicating actions as commanded.

Finally, human soldiers have an unusual responsibility: they are required to disobey illegal orders⁹ (Department of the Army 1956: 183). To do this they must know what illegal orders are, and have the courage to disobey such. They must be able to entertain inconsistent goals: to follow orders, and yet to disobey when those orders are illegal. This requires a judgement based on the moral questions of a specific situation. Without their own values, robots have no basis for making such a judgement. An old publication, *Military Leadership* (Department of the Army 1983: 105, emphasis added), put it this way:

If you ever decide to violate a regulation or law . . . you must be able to show that the situation was so *unique* that you had to violate a regulation, law, or ethical rule to do the most moral thing in terms of the ideals of this nation.

To decide that a situation was unique in that sense requires human intuition, far from a program's abilities.

Military Honor

Just what is involved in military honor? The textbook answer is honesty and courage. We call it the textbook answer because of the material emphasized in the classic, *The Army* (Department of the Army 1981). In the section on the professional army ethic (Department of the Army 1981: 26), the emphasis is on candor (truthfulness) and courage.

More detail is called for. It is the viewpoint of the authors that, other than instrumental values, all other values are created by sacrifice or the risk of sacrifice. A detailed argument for this is found, for example, in Axinn (2009). This position denies value realism, and holds that no values exist independent of humans. Since honor is a value, it

also is created by sacrifice or the risk of sacrifice, and only by that. Honor is a human value, not a property of machines, so its creation requires *that humans risk sacrifice*. Where there is no human in the loop, there is no one to risk sacrifice, and therefore no honor produced. Therefore any nation that deploys autonomous lethal robotic weapons acts without honor. Of course, that nation may risk retaliation in the war at large; however, the actual use of lethal force is done without honor. The commander who launches an autonomous robot is removed from the firing decision and just like the politician who decided to go to war, they take no direct risk of sacrifice from that lethality decision.

Consider a case in which someone says: 'I'll hold your coat while you go and fight.' This is cowardly, not honorable, because the volunteer to hold the coat is detached from the conflict and displays no courage. While some may say that unmanned drone operators display no courage, as there is limited physical risk to them in their control room, they are still deeply involved in the hostilities mentally and they should understand the grave consequences that their tele-operated actions can have. The operator must still have the courage to fire the weapons based on the assumption that the target is legitimate. In the use of a robot operated by an on-board computer, no soldier exhibits courage. This is an advantage toward saving lives on one side, but hardly an example of honor and courage.

Why such concern with honor and the demands of humanitarian behavior? Beyond the inherent moral reasons, the object of war is to make peace with the enemy, peace on desirable terms. If a nation behaves dishonorably, by ignoring the laws of warfare or simple humanitarian matters, their enemy may hate them so much that peace cannot be arranged for a very long time. Therefore, honorable behavior is a useful war strategy, as well as a moral requirement. As noted above, in the language of FM 27-10 (Department of the Army 1956: 3), the purposes of the war conventions include 'Facilitating the restoration of peace.' While there certainly are legitimate 'ruses of war' (as pointed out by the Hague Rules, article 24, and Department of the Army 1956: 22), dishonorable action must be ruled out.

Healthy Emotions

Arguments have been offered holding that robots/drones may do a better job than humans in making target decisions because they have no revenge motive (Arkin 2010). They are not enraged, as humans may be, by the killing of their buddies. But having no emotions, they do not have the attitude toward people that 'healthy' humans are expected to have (Human Rights Watch 2012). They do not realize the enormity of an error in killing the 'wrong' person.

Why is killing with emotions morally superior to killing without emotions? As noted above, honor requires the willingness to risk sacrifice, which in turn requires intention and feeling – emotion. Since morality requires respect for duty rather than following selfish goals, that respect also requires a certain intention or feeling/emotion. The need for human emotion in warfare is cited by at least some authors writing about drones in the general press. For example, John Sifton (2012: 15) says: 'The unique technology allows the mundane and regular violence of military force to be separated further from human emotion. Drones foreshadow the idea that brutality could become detached from humanity – and yield violence that is, as it were, unconscious.'

Furthermore, healthy humans may extend mercy when appropriate. Maj. Daniel L. Davis (2007) has pointed out that:

...in virtually every war involving the US...the enemy discovered that although GIs could be as ruthless and vicious as any opponent, the same soldier could extend mercy when appropriate. As information about US soldiers' humanity spread among enemy combatants, more of them willingly surrendered instead of choosing to continue to fight – which ultimately supports US war aims and saves lives on both sides of the battle line.

The decision on granting mercy requires human emotion. And, importantly, it is essential for the enemy to understand that both sides are governed by human emotions. An enemy is less likely to peacefully surrender to a killer robot and therefore a well-trained human with healthy emotions is more desirable than an autonomous robot for such situations.

Proposal

We propose that autonomous drones be put in the same category that we now use for chemical/gas warfare (and other weapons that are banned). In view of the inhuman features of autonomous robots, we propose that all nations renounce the use of them.

There are various arguments against our proposal, including the fact that similar weapons are now a customary tool of war, since many nations have unmanned drones and have used them. Our proposal is for *autonomous* robots that have lethal ability, not unmanned drones. However, the Ottawa Convention of 1997 dealing with anti-personnel mines was in the same situation, and still managed to ban such weapons. That convention agreed to completely ban a weapon that was in widespread use at the time. In this case, automatic drones are not yet in use, much less in widespread use. As stated in note 6, while most nations of the world have agreed or signed the Anti-Personnel Mine Convention, several key nations have not, including the United States. This does not mean that it is a bad model to follow; rather, that its work is not yet complete.

Defensive Weapons

The use of autonomous robots in a defensive situation, such as securing a perimeter, is a more subtle situation. Conventional perimeter and area denial weapons such as electric fences are allowed due to the fact that any human encountering such a weapon must necessarily have committed an act of hostility. Systems such as autonomous turrets or perimeter patrol robots are thus assumed to be a natural extension of existing weapons systems and are allowed, so long as the perimeter is well marked and anyone, civilian or military, would have adequate warning to avoid the active area.

Defensive weapons are the only type of autonomous *robots* that have been deployed. The United States and other nations use sentry guns to try and shoot down incoming projectiles attacking a ship or other asset (Phalanx CIWS), although they are not designed to attack humans. South Korea has unveiled its SGR-1, which is an autonomous turret that can use thermal imaging to lock on to human-sized targets up to 3 km away for

use in the demilitarized zone (Human Rights Watch 2012). These defensive weapons can only fire on targets that are encroaching a well-delimited area and can therefore be seen as extensions of electric fences.

Support for Our Proposal

Our proposal does have some support in the literature, although most authors do not offer such a specific proposal. There are numerous publications, many of which are cited above, that agree with our conclusion, but for different reasons (e.g. Sharkey 2010, Human Rights Watch 2012). The secondary issues (including discrimination ability and responsibility questions) discussed in the autonomous weapons section are enough to convince some that these weapons should be banned. Others support our views, but do not go so far as to say that they should be banned, stating only that 'we should be wary of the use of robots' (Coker 2008: 151).

Maj. Daniel L. Davis (2007) comes closest in his paper 'Who decides: Man or machine?' where he argues against the use of what he calls 'Mindless Killing Machines,' although he does not go into depth as to why they are immoral or offer a solution. He quotes Col. Lee Fetterman as saying: 'Men should decide to kill other men, not machines. This is a moral imperative that we ignore at great peril to our humanity' (Davis 2007: 3). Our proposal supports and expands on these ideas.

Further and clear support for our conclusions are found in Armin Krishnan's 'Killer robots: Legality and ethicality of autonomous weapons,' and recently in Human Rights Watch (2012) 'Losing Humanity: The Case against Killer Robots.'

Conclusion

For the reasons presented above, reasons both military and humanitarian, we propose that autonomous robots, carrying lethal weapons and operated by computer programs alone, be treated on the same basis as the United States now treats chemical weapons (gas warfare among them). As noted earlier, the United States and all other nations should agree to not use such weapons. A second Ottawa-style conference would be an appropriate mechanism to accomplish this moral goal internationally.

Nuclear weapons are an example of technology that was brought into use before civilization and the laws of war could react to them. We need to act now to establish the moral and legal standing of automatic robots before they enter into common usage. The Ottawa conference prohibiting anti-personnel mines (August 2007), as mentioned, would be an excellent model for what should be done about automatic robots.

A nation that relies on such weapons ignores the humanitarian basis for the laws of war, and an international convention must ban them. As technology continues to progress there will certainly be borderline questions,¹⁰ but the central notion cannot be abandoned: that a lethality decision is to be made only by a human and not a machine. That should remain the key focus of debate and be the guiding moral principle.

ACKNOWLEDGEMENTS

We owe special thanks to two reviewers, anonymous to us, who made very helpful comments on an earlier version of this paper.

NOTES

1. The Geneva Conventions, signed 12 August 1949 in Geneva, Switzerland.
2. A more complete treatment can be found, for example, in Axinn (2009).
3. The Hague Convention respecting the Laws and Customs of War on Land and its annex, signed 18 October 1907 in The Hague, Netherlands. A part of this is known as the Hague Rules.
4. In this paper the references are often to materials from the United States. While we intend the viewpoint of the paper to be international and we propose international concern and action, the United States is currently the largest military (by expenditure) and the largest user of drones. Furthermore, this paper is not intended as a survey of the policies of every nation. Therefore, it is convenient to use mostly US references in the analysis.
5. Convention on Certain Conventional Weapons, signed 10 April 1981, known as CCW.
6. The Convention on the Prohibition of the Use, Stockpiling, Production and Transfer of Anti-Personnel Mines and on their Destruction, signed 3 December 1997 in Ottawa, Canada. As of 15 August 2007, 155 states, more than three quarters of all countries, had agreed or signed, although the United States, Russia, China, and 36 other states are not party to this convention.
7. Nothing in the written or customary laws of war refers to the size of the military power or equipment of each side. Furthermore, the unmanned or autonomous nature of the technology does not change the other just war requirements of proportionality, necessity, just cause, and so on.
8. For a human chess champion to lose to a computer may reshape how we think about technology, but it raises no moral problems as a chess loss is ordinarily not lethal.
9. Also from article 92 of the Uniform Code of Military Justice (UCMJ), 10 USC Chapter 47.
10. Although we have given a general argument, there will always be edge cases and refinements. In unmanned robots, the following issues may arise: what if a human is controlling two drones? What if they are controlling 10 drones? One hundred drones? What if the drone has an automated target suggestion routine that still prompts the operator to approve possible targets? What if the operators start to rely on the machine's judgement over their own?

REFERENCES

- Arkin, Ronald C. (2009) *Governing Lethal Behavior in Autonomous Robots* (Boca Raton, FL: CRC Press).
- Arkin, Ronald C. (2010) The Case for Ethical Autonomy in Unmanned Systems, *Journal of Military Ethics*, 9(4), pp. 332–341.
- Asimov, Isaac (1950) *I-Robot* (New York: Doubleday and Company).
- Axinn, Sidney (2009) *A Moral Military, Revised and Expanded Edition* (Philadelphia, PA: Temple University Press).
- Axinn, Sidney (2010) *Sacrifice and Value: A Kantian Interpretation* (Lanham, MD: Lexington Books).
- Burroughs, John (1999) The Advisory Opinion of the International Court of Justice and the Statute of the International Criminal Court, paper presented at a conference on 'Nuclear Policy and Security on the Eve of the 21st Century,' St. Petersburg, Russia, 18–20 June 1999, accessed 15 July 2013, available at: <http://www.lcnp.org/wcourt/humlaw.htm>; Internet.
- Coker, Christopher (2008) *Ethics and War in the 21st Century* (New York: Routledge).

- Davis, Maj. Daniel L. (2007) 'Who Decides: Man or Machine?', *Armed Forces Journal*, accessed 1 July 2013, available at: <http://www.armedforcesjournal.com/2007/11/3036753>; Internet.
- Department of the Army (1956) *Department of the Army Field Manual FM 27-10: The Law of Land Warfare*, No. 1 (July 1976), accessed 1 July 2013, available at: http://www.loc.gov/rr/frd/Military_Law/pdf/law_warfare-1956.pdf; Internet. [Still authoritative].
- Department of the Army (1981) *Department of the Army Field Manual FM 100-1 The Army* (Washington, DC: US Government Printing Office).
- Department of the Army (1983) *Department of the Army Field Manual FM 22-100: Military Leadership* (Washington, DC: US Government Printing Office).
- Human Rights Watch (2012) *Losing Humanity: The Case against Killer Robots*, accessed 25 May 2013, available at: <http://www.hrw.org/reports/2012/11/19/losing-humanity-0>; Internet.
- Kant, Immanuel [1785] (1959) *Foundations of the Metaphysics of Morals*, Lewis White Beck (trans.) (New York: The Liberal Arts Press).
- Kant, Immanuel [1787 2nd edition] (1965) *Critique of Pure Reason*, trans. Norman Kemp Smith (New York: St. Martin's Press).
- Kucinich, Dennis (2012) Combat Drone Programs, *Congressional Record*, 158(146), p. H6377.
- Krishnan, Armin (2009) *Killer Robot's Legality and Ethicality of Autonomous Weapons* (Farnham, England: Ashgate Publishers Ltd).
- Ministry of Defence (2011) *JD 2/11: The UK Approach to Unmanned Aircraft Systems*, accessed 1 July 2013, available at: <https://www.gov.uk/government/publications/jdn-2-11-the-uk-approach-to-unmanned-aircraft-systems>; Internet.
- Sifton, John (2012) A Brief History of Drones, *The Nation*, 27 February, accessed 1 July 2013, available at: <http://www.thenation.com/article/166124/brief-history-drones>; Internet.
- Singer, Peter W. (2009) *Wired For War: The Robotics Revolution and Conflict in the Twenty-First Century* (New York: The Penguin Press).
- Singer, Peter W. (2012) Do Drones Undermine Democracy? *The New York Times*, 22 January, accessed 25 May 2013, available at: <http://www.nytimes.com/2012/01/22/opinion/sunday/do-drones-undermine-democracy.html?pagewanted=all&r=0>; Internet.
- Sharkey, Noel (2008) Grounds for Discrimination: Autonomous Robot Weapons, *RUSI Defense Systems*, 11(2), pp. 86–89.
- Sharkey, Noel (2010) Saying 'NO!' to Lethal Autonomous Targeting, *Journal of Military Ethics*, 9(4), pp. 369–383.
- Sparrow, Robert (2007) Killer Robots, *Journal of Applied Philosophy*, 24(1), pp. 62–77.
- Strawser, Bradley J. (2010) Moral Predators: The Duty to Employ Uninhabited Aerial Vehicles, *Journal of Military Ethics*, 9(4), pp. 342–368.
- Szilard, Leo (1945) *A Petition to the President of the United States*, 17 July, accessed 1 July 2013, available at: <http://www.dannen.com/decision/45-07-17.html>; Internet.

Aaron M. Johnson is a PhD candidate in the GRASP Laboratory, Electrical and Engineering Department, University of Pennsylvania. His research interest is in novel robotic design – behavioral design, controller design, and platform design – and he has worked on a wide variety of non-weaponized robots for military and civilian use, including snake robots, legged robots, and the iRobot Pack Bot. Correspondence address: Electrical and Systems Engineering Dept., 200 S. 33rd St., Moore building 203, Philadelphia, PA 19104, USA. E-mail: aaronjohnson@alumni.cmu.edu

Sidney Axinn is currently Courtesy Professor in the Philosophy Department at the University of South Florida. He chaired the Philosophy Department at Temple University for 15 years, and is Professor Emeritus there. He has taught or lectured in Japan, South Korea, Hungary, England, and Belgium, as well as West Point and the South Korea Military Academy. He is the author of *A Moral Military – Revised Edition*, *Sacrifice and Value: A Kantian Interpretation*, *The Logic of Hope: Extensions of Kant's View of Religion*, and over 40 articles in philosophy journals. A Second World War veteran, he served with the US Army in North Africa, Italy, and Corsica (in the 1073 Signal Co. attached to the 12th Air Force). He is also the grandfather of the first author. Correspondence address: 3300 Darby Road, Apt. #2207, Haverford, PA, 19041, USA. E-mail: axinn@temple.edu

Copyright of Journal of Military Ethics is the property of Routledge and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.