***The Ethics of Autonomous Weaponry; A philosophical analysis***

**Name: Orlando Mednick**

**ID: i6046036**

**Tutor: Maarten Verkerk**

**Course: Philosophical Ethics**

**Word Count: 4000**

***Introduction***

Autonomous weapons are currently in development by the US military. They would enable governments to send drones into enemy countries entirely unmanned and autonomous. They would theoretically be able to locate and kill a target without human control (Department of Defense, 2012). Furthermore, they are supposed to be able to work in groups, and be fully aware of their surroundings. If this technology comes into active development, it would lead to a change of warfare as we know it (Brooker, 2013). It would significantly reduce the casualties within war, but it would also reduce the amount of accountability within the army’s hierarchy. It would also only reduce the amount of casualties on the side with the drones, therefore creating heavily asymmetrical warfare. The ethics of autonomous weaponry is something that must be taken into consideration before they can be used in the field however. Should we impose a pre-emptive ban on autonomous weapons? Can we use classical philosophy to help us make a decision? Firstly I will outline the main arguments for and against the use of autonomous weaponry, and then analyse the issue; applying classical philosophical thinking with the aim of judging the morality of using autonomous weapons. I will be using Kant’s Deontology, Utilitarianism and Levinas’ ideas on the role of the face in moral decisions. I use the term drone to describe autonomous weapons, I will use the term unmanned drone to refer to those that are controlled directly by humans.

***The dangers of autonomous weaponry***

The current development of these autonomous weapons is currently a very controversial topic in the media. There is a notable “Stop the Killer Robots” campaign starting in April in which notable academics, Nobel Peace Prize laureates, will take the issue to the House of Commons and attempt to create a pre-emptive ban on autonomous weapons (McVeigh, 2013). Autonomous weapons is an idea that has been played with by Science Fiction for decades, but the reality would not resemble the fiction. The robots would not be able to think as humans do, and not make rational and moral judgements (Sharkey, 2012). The ethical debate about how these can be used is extremely controversial and the US government are currently drafting bills on these weapons. But what are the problems with autonomous weapons? These are the main arguments against the use of autonomous weapons in warfare, and they outline the moral problems with autonomous weapons.

 The first major problem with autonomous weapons is that it can potentially reduce the accountability within the army. If an autonomous drone malfunctions and kills a group of civilians, who is to blame? You can blame anyone from the officer who allowed the drone to go into the field, to the man who wrote the robot’s code (Sharkey, 2012). This means that if there are negative consequences following deployment of a drone, there is no one to hold fully accountable. A lack of accountability in any situation is bad, but during wartime it is worse. This is because one must be able to be held accountable for misdeeds if the integrity of the army is to be maintained (Friedkin, 2012). The modern worlds justice system is founded on the notion of accountability, and the remove this from the global community would mean a step back from global unity. Furthermore, if one is not held accountable, or the wrong person held accountable, due to misunderstanding responsibility, then there is potential that the action of the drone was deliberately planned, and the perpetrator is unpunished. Although there would be human oversight at a top level, they would not necessarily be responsible for the action of the drone if it malfunctioned.

 The second major problem with autonomous weaponry is that there is no current guarantee that they will be able to distinguish between target and civilian. The technology to create this distinction is not currently available, and is not seen to be realistically achievable by the time these autonomous drones are available (Sharkey, 2012). This means that there can potentially be drones that would not avoid harming civilians in order to take down a target or, worse still, could mistake a target for a civilian. Furthermore they would not be able to make the rational judgements that humans can make in order to minimize or avoid civilian casualties, for example delaying the elimination of a threat due to the proximity to a local school (Sharkey, The Automation and Proliferation of Military Drones and the Protection of Civilians , 2013). Artificial intelligence is currently extremely far away from giving human judgments and emotions, however it is easy for people to make the mistake of attributing human emotions to machines. This is because of the influence science fiction has had upon the modern view of robotics; the media presents us with constant images of human like robots with emotions similar to our own (Sharkey, The evitability of autonomous robot warfare, 2012). It would be impossible for scientists to develop a computer as complex as the human brain, thus emotions of a machine would be extremely limited. They would not be familiar emotions to you or me, but due to the influence of the media, it is could be easy for us to interpret them as such.

 Another of the major problems with autonomous weaponry is that it heavily affects symmetrical warfare. This is an ethical problem because it means that poorer countries who cannot afford this weaponry will have a drastically higher amount of casualties than the country possessing the weaponry. This argument is never addressed in the development of weapons for obvious reasons, but it poses a great ethical issue with modern warfare. The development of drones and other unmanned weapons is an issue to begin with, but with autonomous weapons there would be no human control at all, thus a far higher deployment rate would be available (International Committee for Robot Arms Control, 2009). If we apply this to reality we can see the disproportionality; if there was a war in Africa that America got involved in if America was in possession of autonomous weapons, it would be ill equipped African militia, most commonly using the AK47 against autonomous drones which would be able to hunt and kill targets with the latest technology. This disproportionality causes a lot of difficulties within the battlefield, as it would also influence other countries into developing heavier and more deadly weapons. Rivalry within weapons development will reoccur if new forms of weaponry are developed, thus paving the way for a far more dangerous world.

***Advantages of autonomous weaponry***

Even though there are problems with autonomous weaponry, there are advantages in the use of them as well. The US government have been drafting bills to regulate the use of these potential weapons, and make sure that there is enough oversight from human authority. These new drones would be able to be activated or deactivated at any point if it comes to it and they would undergo rigorous checks to make sure they are in fully functional condition. These regulations make sure that the drones would not be a threat when it comes to malfunctioning, they would simply be able to be deactivated (Mehta, 2012). This adds a level of accountability to the situation, but it is still not enough for adequate insurance of genuine accountability. The fact that the US has already taken these factors into consideration shows that there is careful deliberation in the implementation of them in the battlefield, and ethical arguments are being brought into question. The main factor taken into account is human oversight; it means that judgements can be made by humans and the drones would be able to interpret this. The final decision is made by a commanding figure, and the drones themselves will need approval undersecretary of defence for policy, the undersecretary of defence for acquisition, technology and logistics, and the chairman of the Joint Chiefs of Staff before their activation (Mehta, 2012).

 Another advantage of autonomous weapons is that it can also be a counter to current asymmetrical warfare. This argument can be used both for and against autonomous weapons, as the current global arms situation is extremely complex. China and Iran’s current cyber weaponry and drone technology is very advanced, and the development of these drones would help to equalise the difference in technology (Thurnher, 2013); however the case stated in the last section also holds true, the majority of countries are far behind this level of technology in weaponry. This does not mean that we should simply stop developing technology so the rest of the world can catch up however, it simply shows the complexity of symmetrical warfare. If the other superpowers in the world are becoming more advanced with their weaponry, it seems appropriate to arm ourselves further to increase the security of your own country. Security is an extremely important issue in today’s world, the inflammatory nature of inter-state relationships shows how important it is to protect the citizens. For the USA it is especially important as they have tender relationships with other major powers. If China or Iran developed weapons at a much faster rate than western countries it would be possible for there to be increased risk of war or antagonism. Superpowers must remain on top of technological research and the latest weaponry to insure their safety.

 The final argument I will outline for the use of autonomous weaponry is that artificial intelligence is developing at an incredible rate. The current Google cars can learn from previous mistakes, and can drive to locations entirely autonomously (Carlson, 2013). This idea can be applied to drones, especially if they can work as teams. This means that even if the drones make “mistakes” they can learn from this and change future behaviour, which increases the reliability of the drones by allowing improvement within its own software. The autonomous aspect of the drones would also be able to be used in many other parts of the military that are not just “killer robots”. The development of aircrafts that would be able to fly routes just by being given a flight path would make the Air Force far more useful as they would need less human interaction for control (Mehta, 2012). Drones would also be used for infiltration and investigation; having unmanned spy drones staying in locations for days or weeks, following targets and mapping areas (Brooker, 2013). This would revolutionise the collection of intelligence, and make a far more reliable and effective means to end conflict. Furthermore, if these weapons can identify targets as well as normal soldiers in the future, it would substantially reduce the amount of casualties suffered during wartime and could insure a far more humane method of dealing with hostile targets.

***Deontology and autonomous weaponry***

To get a full understanding of how to approach the ethical issue of autonomous weaponry we can apply classical philosophical thinking to the issue at hand. Kant’s deontology can give insight into how to make an ethical judgement. Kant proposed the categorical imperative as a method of making rational moral decisions. He claimed that we must apply the categorical imperative to every moral decision we make, we need to see if the action can be universalised, and if it can, then it is our duty to act in that way. This means that killing is never allowed under deontology, as you can never universalise killing (Kant, 1998). The situation of war can never be justified with strict deontology, Kant disapproved of war and believed that there was no morality within leaders who are involved in military conflict. This means that autonomous weaponry would not be allowed under the categorical imperative. There can be some justification however, in the use of drones if we interpret deontology loosely. This is because the use of drones means that we are not using people as a means to an end (Rudmin, 2010). Autonomous drone warfare would reduce the amount of soldiers needed in the battlefield, thus the government would not be using the people as a means to an end anymore, and they would be using autonomous drones to fulfil their objective. This is not an entirely strong case though, the use of drones would reduce the amount of soldiers in the battlefield, but the killing could not be universalised, and there would not be any killing until both sides fought entirely with autonomous weaponry; if it got to this point then the war would be trivial.

 Drone usage in warfare would also result in distorted reciprocity, and the universal maxim would not be exercised. This creates an issue with deontology as every moral action must conform to the categorical imperative. We can look to a hypothetical example of this distortion of reciprocity if we reverse the situation and imagine that Libya had sent drones to America (Fetzer, 2011). The response to this from America would be devastating, and they would respond with mass destruction. This shows unequal acceptance of attacks, and thus does not fit into the universalisation maxim because the drone use is not equally accepted universally. We can see the principle of not using humans as a means to an end as a criticism also, as proponents of drone warfare see civilian casualties as “collateral damage” to justify the resolution of a just war (Journal of Faith and War, 2012). This does not fit with deontology, as civilian casualties are never accepted within the categorical imperative and humans are being used as means to an end. Deontology simply is not compatible with the use of autonomous weapons or any sort of warfare. Warfare will always use people as a means to an end, regardless of whether autonomous weapons are involved, or the war is declared a just war. Deontology can never condone the use of machines or actions used for political gain which would harm humans.

***Utilitarianism and autonomous weaponry***

When applying utilitarianism to the idea of autonomous weapons we must first try and predict what the consequence of their use would be, and then frame our ethical ideas from there. The main principle of utility proposed by Jeremy Bentham is to maximise happiness for the maximum number of people (Bentham, 1996). This means that if the use of autonomous weaponry would hasten the resolution of the conflict, and if their use would minimize civilian and soldier casualties, then the use of such weaponry would be acceptable. The use of autonomous weaponry would significantly reduce the amount of casualties sustained during wartime due to their potential for replacing human soldiers. This would be the first element of happiness that would be maximised, as there would be the happiness of the soldier being alive still, the happiness of their family and overall societal morale. This means that autonomous weaponry can potentially be an ethical alternative during wartime (Littlejohn, 2013). The element of asymmetrical warfare does not enter into this equation as the casualties sustained on the opposing side of the forces would theoretically be either equal or less than if human soldiers were used. Furthermore, if the use of this technology becomes more readily available, then the issue of symmetrical warfare may be overcome and the happiness would be increased on both sides, due to the decrease in soldiers used in the military on both forces. If the intelligence gathered by the drones brought the end of the conflict quicker it would be further reason to use drones in conflict, as they could be used for all purposes, not just for target elimination.

 There are issues with using utilitarianism to evaluate the ethics behind autonomous weaponry however. Firstly it is naïve to think that we can predict the consequences of the actions of these drones in conflict, especially since they are not currently existent. To make presumptions about how they will alter the face of conflict before they have been developed is dangerous, and unpredictable events could happen. This is the issue with applying consequentialist approaches to ethics, especially with topics that could easily have unforeseen consequences, for better or for worse. But by the same token, it would also be unwise to ban the weapons pre-emptively due to the same idea. We cannot see what the consequences will for certain, thus we must wait to have first-hand evidence to support a claim either way, but then the issue could be that it would then already be too late. Utilitarianism lacks strength in its arguments due to these reasons, we can’t see the consequences thus we must wait until the technology already exists to judge it. Consequentialism implies that we can only make ethical judgements once we know the entire situation, and the consequences, thus it is in essence a retrospective ethical theory. This creates problems when trying to judge the ethics of a prospective technology; thus makes utilitarianism insufficient when basing an ethical argument for autonomous weaponry. If one believes that we have enough information to make a judgement such as this, then utilitarianism would agree with the use of autonomous weaponry.

***Levinas and autonomous weaponry***

War throughout time has predominantly been conflict between two ideological factions of humans. Thus war and violence is in essence just violent human interaction. Levinas’ theories put the “other” (the person one is interacting with) as the supreme authority in that interaction. To label Levinas’ ideas as an ethical theory is slightly wrong, instead it is more of an analysis of human interaction, and a basis for how this interaction determines relationships. To Levinas, to see the other’s face is the most important factor in an ethical interaction. This is to look past features such as hair colour, eye colour or skin colour and to look at the being they are (Levinas, 1995). This gives us an initial critical view on war, as war is waged without seeing the other as a person, simply looking at features of the body or personality. There is no interaction between the people in charge, and if there is then they fail to hold the other as the supreme authority, as if both did that then there would be no conflict. To respect the other as higher than you means that you have upmost respect for every individual. Autonomous weapons go further than removing the interaction between soldiers in the battlefield, it makes the value of the other’s life even lower than in usual war due to the detachment there is when killing them (Ivry, 2010). There is no interaction between me and the other with autonomous weaponry, it is interaction between a drone whose only objective is to kill, and the other, who most definitely does not hold the higher authority.

 The proximity between me and the other is extremely important to Levinas, as he believed if you face the other and look at them as a person, then you would not be able to kill them (Levinas, 1995). Autonomous weaponry removes the proximity as well as the interaction, the one who orders the drone would be thousands of miles away from the target. This creates ethical issues as it removes the interaction that is necessary within all human relationships, whether violent or not. There is no easy sign for peace if neither side is looking to look at the other as a person, nor there is definitely no idea of holding the other as the higher authority if one is sending entirely disposable machinery to kill. There is no chance that Levinas would approve of the use of autonomous weapons. The fact that it removes human interaction would be an initial argument following the basic argument against war in general. Levinas kept the idea of the face extremely important, and the superiority of the other is central to his philosophy, and the exclusion of the other is entirely against what he spoke of (Ivry, 2010). There would be no ethical argument for the use of autonomous weapons from Levinas, and in fact it would make human lives seem even more disposable, reducing the value of the other even further.

***My thoughts***

The development of autonomous weaponry simply seems to show what the global attitude towards human life has become. The value of life to leaders of militaries seems to be so little that they seek to make robots do their killing for them. Human interaction is important, it shows that there is some element of emotion involved, the soldiers fighting the wars today have determination; they believe they are fighting for their country and are being patriots. Whether or not you agree with a specific war, this is the case of the soldiers giving their lives in the war. To remove these soldiers and instead put in emotionless machinery to kill is unethical. Whether or not these machines are effective in reducing ally casualties, it shows that huge spending and top tier research is being given to discovering new ways to kill. The ethics of the weapons themselves aren’t the main issue, it is what it reveals about human nature and those in control of the world. The development of autonomous killer robots such as these shows how little regard leaders have for civilian casualties and human life. A pre-emptive ban on this weaponry is wise in my opinion, there is far too much as risk for international relations and the ethics of warfare. The competitive nature of the international community is also means for a ban on this weaponry, as competition of weapons escalates quickly. If weapons are developed by the USA such as these, then other superpowers will feel the need to develop even more effective killing technology, and thus the cycle continues. This would imply that autonomous weapons would only be the beginning, and much more sinister means for destruction would follow. If there is an international ban on these weapons it would show that there is some sort of international justice and ethics.

 I also believe that it is sad to see that weapons is always the first call of action when it comes to latest technology. If the same funding and technology went into developing tools for humanitarian aid, the world could be substantially improved. For example autonomous devices could be developed that would supply food to those in impoverished countries. There could be food drops for refugees which do not need pilots, they would simply be able to seek their own flight path. These drones could potentially even be used for domestic help, such as garbage collection or little picking. Although these domestic examples may seem trivial, it is simply proposals of alternatives to weaponry that could be easily implemented. I believe that using autonomous weapons to kill trivialises human life and makes the value of the individual in question entirely minimal. Diminishing the value of human life is clearly unethical, and it creates major problems in warfare. I feel that autonomous weapons would pave the way to a new kind of warfare, one which is much more sinister and deadly than we currently possess. If no soldiers’ lives are being lost, then more drones can be deployed, as one wouldn’t have to worry about casualties, therefore the opposition would face an even larger onslaught of attack.

***Conclusion***

Autonomous weapons are potentially extremely dangerous. The arguments in support of autonomous weapons are weak, and they give no ethical reason for the weapons to exist, only logical reasons. There are many things in the world which can easily be interpreted as logical, but are definitely not ethical. The philosophical arguments show that there is no ethical support for autonomous weapons, with the exception of utilitarianism that does not give a conclusive argument either way. For governments to develop these weapons show how little regard they have for the lives of those in foreign countries. The detachment from the conflict and simply sending in robots to do your dirty work for you is unacceptable and should not be the way the world is heading. Deontology would strictly forbid the use of these weapons, as it would not be able to be universalised and it uses humans as a means to an end, which is a valid argument against the use of these weapons, although it is difficult to find any act that can be universalised under the categorical imperative. Levinas would definitely disapprove of the use of these weapons, since it detaches human interaction and diminishes the role of the other to the extent that there is no longer any respect for their life, not giving any sort of human interaction.

***References***

Bentham, J. (1996). *An Introduction to the Principles of Morals and Legislation.* Oxford: Oxford University Press.

Brooker, C. (2013, February 13). *I know in my bones that a robot is going to kill you – the new micro-drones*. Retrieved from The Guardian: http://www.guardian.co.uk/commentisfree/2013/feb/24/new-wave-of-micro-drones

Carlson, N. (2013, March 11). *Google Is Working On A Technology That, If Perfected, Would Save 1.2 Million Lives Per Year*. Retrieved from Business Insider: http://www.businessinsider.com/google-technology-saving-12-million-lives-2013-3

Department of Defense. (2012). *Autonomy in Weapon Systems.* USA: Department of Defense.

Fetzer, J. (2011, February 22). *On the Ethical Conduct of Warfare: Predator Drones*. Retrieved from Global Research: http://www.globalresearch.ca/on-the-ethical-conduct-of-warfare-predator-drones/23324

Friedkin, Z. (2012, November 30). *The Kantian Case Against Drone Warfare*. Retrieved from Big Think: http://bigthink.com/praxis/remote-controlled-morality

International Committee for Robot Arms Control. (2009, September). *Mission Statement*. Retrieved from International Committee for Robot Arms Control: http://icrac.net/statements/

Ivry, B. (2010, February 19). *A Loving Levinas on War*. Retrieved from The Jewish Daily Forward: http://forward.com/articles/125385/a-loving-levinas-on-war/

Journal of Faith and War. (2012, July 2). *The Moral Crisis of Just War: Beyond Deontology toward a Professional Military Ethic - The Crisis of the Deontological Vision of Just War* . Retrieved from Faith and War: http://faithandwar.org/index.php?option=com\_content&view=article&id=158%3Athe-moral-crisis-of-just-war&catid=43%3Ahistory-of-war&Itemid=58&limitstart=2

Kant, I. (1998). Groundwork for the Metaphysic of Morals. In M. Gregor, *Groundwork for the Metaphysic of Morals* (pp. 2-3, 31-39). Cambridge: Cambridge University Press.

Levinas, E. (1995). *Ethics and infinity. Conversations with Philippe Nemo.* Pittsburgh: Duquesne University Press.

Littlejohn, B. (2013, February 20). *Drones, Prudence, and Pre-Emption*. Retrieved from There is Power in the Blog: http://www.politicaltheology.com/blog/drones-prudence-and/

McVeigh, T. (2013, February 2013). *Killer robots must be stopped, say campaigners*. Retrieved from The Guardian: http://www.guardian.co.uk/technology/2013/feb/23/stop-killer-robots

Mehta, A. (2012, November 27). *U.S. DoD’s Autonomous Weapons Directive Keeps Man in the Loop*. Retrieved from Defense News: http://www.defensenews.com/article/20121127/DEFREG02/311270005/U-S-DoD-8217-s-Autonomous-Weapons-Directive-Keeps-Man-Loop

Rudmin, F. (2010, January 1). *Kant on War*. Retrieved from CounterPunch: http://www.counterpunch.org/2010/01/01/kant-on-war/

Sharkey, N. (2012). The evitability of autonomous robot warfare. *International Review of the Red Cross*.

Sharkey, N. (2013, March 28). The Automation and Proliferation of Military Drones and the Protection of Civilians . Sheffield, United Kingdom.

Thurnher, J. (2013, March 28). *No One at the Controls: Legal Implications of Fully Autonomous Targeting*. Retrieved from NDU Press: http://www.ndu.edu/press/fully-autonomous-targeting.html