



This is a repository copy of *Autonomous weapons systems, killer robots and human dignity*.

White Rose Research Online URL for this paper:
<https://eprints.whiterose.ac.uk/139775/>

Version: Published Version

Article:

Sharkey, A.J. orcid.org/0000-0002-4713-2575 (2019) *Autonomous weapons systems, killer robots and human dignity*. *Ethics and Information Technology*, 21 (2). pp. 75-87. ISSN 1388-1957

<https://doi.org/10.1007/s10676-018-9494-0>

Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:
<https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>



Autonomous weapons systems, killer robots and human dignity

Amanda Sharkey¹

© The Author(s) 2018

Abstract

One of the several reasons given in calls for the prohibition of autonomous weapons systems (AWS) is that they are against human dignity (Asaro in *Int Rev Red Cross* 94(886):687–709, 2012; Docherty in *Shaking the foundations: the human rights implications of killer robots*, Human Rights Watch, New York, 2014; Heyns in *S Afr J Hum Rights* 33(1):46–71, 2017; Ulgen in *Human dignity in an age of autonomous weapons: are we in danger of losing an ‘elementary consideration of humanity’?* 2016). However there have been criticisms of the reliance on human dignity in arguments against AWS (Birnbacher in *Autonomous weapons systems: law, ethics, policy*, Cambridge University Press, Cambridge, 2016; Pop in *Autonomous weapons systems: a threat to human dignity?* 2018; Saxton in *(Un)dignified killer robots? The problem with the human dignity argument*, 2016). This paper critically examines the relationship between human dignity and AWS. Three main types of objection to AWS are identified; (i) arguments based on technology and the ability of AWS to conform to international humanitarian law; (ii) deontological arguments based on the need for human judgement and meaningful human control, including arguments based on human dignity; (iii) consequentialist reasons about their effects on global stability and the likelihood of going to war. An account is provided of the claims made about human dignity and AWS, of the criticisms of these claims, and of the several meanings of ‘dignity’. It is concluded that although there are several ways in which AWS can be said to be against human dignity, they are not unique in this respect. There are other weapons, and other technologies, that also compromise human dignity. Given this, and the ambiguities inherent in the concept, it is wiser to draw on several types of objections in arguments against AWS, and not to rely exclusively on human dignity.

Keywords Autonomous weapons systems · Human dignity · Killer robots · International humanitarian law · Laws of war · Moral machines · Robot ethics

Robots and computers can be used to make decisions that affect humans in many spheres of life, from the trivially domestic (e.g. when to turn the heating on), to situations of life and death (e.g. autonomous car accidents). Autonomous weapons systems (AWS) represent an extreme example of such decision-making. These are weapon systems that can select and engage targets without human intervention. In other words, weapons systems that can make a decision to take human lives.

There has been considerable discussion and debate over definitions of autonomy in weapons systems, but an overlap and agreement in definitions is beginning to emerge. The US Department of Defence defines AWS as weapons

that are able, ‘once activated, to select and engage targets without further intervention by a human operator’ (DOD directive 2012, updated 2017). The International Committee of the Red Cross (ICRC) defines them as ‘weapons that can independently select and attack targets, i.e. with autonomy in the ‘critical functions’ of acquiring, tracking, selecting and attacking targets’ (ICRC 2014). Recent definitions often also incorporate the notion of meaningful human control: thus Human Rights Watch defines AWS as weapons which ‘would identify and fire on targets without meaningful human control’ (HRW 2014; Amoroso et al. 2018).

Concerted efforts are currently being made to prohibit the development and use of AWS. The ‘campaign to stop killer robots’ was launched in 2013, and is formed of a global coalition of over 72 non-governmental organisations (NGOs) from 31 countries. In 2015, an open letter was released, calling for ‘a ban on offensive autonomous weapons beyond meaningful human control’, and signed by

✉ Amanda Sharkey
a.sharkey@sheffield.ac.uk

¹ Department of Computer Science, University of Sheffield, Sheffield S1 4DP, UK

nearly 4000 Artificial Intelligence and Robotics researchers, and over 22,000 others. There have been discussions at the United Nations of these weapon systems: the state parties to the Convention on Conventional Weapons have held annual meetings since 2014, and a Group of Governmental Experts was established at the UN in 2016 for further discussions, the most recent of which were held in November 2018. Further discussions are being held in March 2019. In 2018, a call to ban these weapons was supported by 26 states.

There are many strong reasons to support a ban of AWS. These range from extreme difficulties in complying with international law to serious problems with global security. It has also been said that they are against human dignity (Asaro 2012; Docherty 2014; Heyns 2017; Ulgen 2016). However there have been criticisms of the use of human dignity as the basis for arguments for a prohibition of AWS (Pop 2018; Saxton 2016; Birnbacher 2016). There is also some uncertainty about the meaning of ‘dignity’ itself, and about its usefulness (e.g. Werner 2014). There is a need for greater clarity on this issue. The aim of this paper is to critically examine the relationship between human dignity and the arguments against AWS.

This paper begins by outlining the main arguments that have been made against autonomous weapons. Then we will look at the claims that these weapons are against human dignity, and their criticisms, framing these in the context of differing accounts of the meaning of human dignity. Following this, four questions about the relationship between human dignity and arguments against killer robots will be identified and discussed.

Against autonomous weapons systems

A 2012 Human Rights Watch report provides an account of a comprehensive set of arguments against AWS (Human Rights Watch 2012). These arguments are based on the likely impact of AWS on the risks faced by civilians in war, and on their inability to conform to international humanitarian law (IHL), also known as the laws of armed conflict, or the laws of war. The laws of war reflect the ‘exceptional circumstances that prevail during armed conflict’ (Heyns 2016, p. 8) where there is a need to protect those not involved in the conflict. The principles of distinction, proportionality and military necessity are the crucial aspects of IHL intended to protect civilians. The 2012 report is clear that AWS are not able to realise these principles, since they require human judgement and human understanding. The report also points out that eliminating human involvement would remove the opportunity for compassion, which can provide a means for reducing the amount of civilian deaths. In addition, the use of AWS could make going to war more likely, since politicians could instigate conflict without risking their nation’s

human soldiers. Another tool for civilian protection identified in the report is accountability. When autonomous weapons are used there is a lack of accountability, since the weapon or robot cannot be held responsible for its actions, and if unnecessary civilian deaths or casualties were to occur it is not clear who could be punished or held to account for them. Related arguments are made by several writers, including Sharkey (2012a, b), Asaro (2012), Tamburrini (2016) and Heyns (2013, 2016, 2017).

There are arguments that focus on the extent to which AWS *can* adhere to IHL and the laws of war, and arguments that focus more on whether they *should* be used, even if they were shown to be capable of doing so. In arguments that correspond to the first category, Sharkey (2012a, b) highlights the limited abilities of programmed computational systems and robots to conform to IHL. For example, the principle of distinction refers to the requirement that a weapons system must allow the discrimination of combatants from non-combatants or other immune actors. Sharkey (2012a, b) argues that AWS lack three necessary components for this. First, although they would be able to detect humans, their sensory and vision systems are not able to reliably tell combatants from non-combatants, or other immune actors such as wounded combatants, or those who have surrendered. Second, there is no existing codified, or programmable, definition of what constitutes a civilian (see also Sharkey 2008). And third, autonomous robots and weapons lack the situational understanding and battlefield awareness that is needed to satisfy the principle of distinction. For instance, a human could draw on an understanding of social situations to recognise insurgents burying their dead, or children being forced to carry rifles, in a way that a robot could not.

The principles of proportionality and military necessity are also beyond the capabilities of present and near future robots and weapons systems. Sharkey (2012a, b) distinguishes between what he calls the ‘easy’ and ‘hard’ proportionality problems. The easy version involves calculating the likely collateral damage of different forms of attack, and directing an attack so as to minimise such damage. For instance, it is conceivable that robot software could choose the munitions to be used near a school in order to minimise the number of children killed. However, the hard version of proportionality is deciding whether the military advantage to be gained in such a situation would justify the use of any form of attack near a school. Decisions about military advantage and military necessity require ‘responsible accountable human commanders, who can weigh the options based on experience and situational awareness’ (Sharkey 2012a, b). Suchman (2016) has also argued that machines cannot fulfil the requirements of situational awareness. Related arguments about the extent to which robots and computational devices could be programmed or trained to develop the necessary moral competence are reviewed in

Sharkey (2017). There are reasons to believe that not only do they not have situational understanding at present, but also that they are unlikely to have such understanding in the foreseeable future.

Heyns (2013) also points out the current inability of autonomous weapons to do proper targeting, and their “lack of human judgment, common sense, appreciation of the larger picture, understanding of the intentions behind people’s actions, and understanding of values and anticipation of the direction in which events are unfolding” (Heyns 2013). However, as well as questioning whether autonomous weapons *can* do proper targeting, he also asks whether they *should* be used even if they were able to adhere to IHL rules about distinction, proportionality and precaution (Heyns 2017).

Heyns’ argument is that even if autonomous weapons were able to match or exceed human ability to conform to IHL, they should still not be used without meaningful human control. According to him, their use would be an offence against the right to life because (i) errors would still be made and there would be no person or persons to be held accountable, and (ii) the lack of human deliberation would make targeting decisions arbitrary. He also argues that they would be against the right to dignity of those targeted and of those in whose name such force was deployed. Heyns (2017) bases his argument on the African Charter on Human and People’s Rights, which is emphatic about the need to protect a right to a ‘dignified life’, and which emphasises the interrelated nature of the right to life and the right to dignity. In other work (Heyns 2013, 2016), he places more emphasis on a Kantian account of dignity.

Asaro (2012) also argues that AWS should not be used even if they were able to meet the requirements of IHL. He argues that the IHL governing armed conflict, and the principles of distinction, proportionality and military necessity, imply a requirement for *human* judgement, and a duty not to delegate the capability to initiate the use of lethal force to unsupervised machines or automated processes. His suggestion is that IHL should be updated to include a prohibition of AWS, and the establishment of the principle that taking a human life requires an informed and considered human judgement.

Asaro’s contention is that IHL requires human judgement because the ‘rules’ that constitute it require interpretation, and are quite unlike the explicit rules of chess. For example, the International Committee of the Red Cross has developed guidelines to determine when a person should be considered to be a combatant. These guidelines specify three requirements that must be satisfied before it can be concluded that a civilian is a legitimate target: (i) their actions must cross a threshold of harm, and adversely affect military operations; (ii) the harm must be directly caused by their actions and (iii) their actions must meet the requirement of *belligerent*

nexus and be designed to directly cause the threshold of harm in support of one party to an armed conflict, to the detriment of another. Guidelines such as these cannot be easily translated into the kind of unambiguous rules that a robot or computer system can follow.

According to Asaro (2012) ‘the very nature of IHL presupposes that combatants will be human agents’ (p. 2). Its rules are supplemented by heuristic guidelines for humans to follow, and it requires combatants to consider the implications of their actions. He also argues that law in general requires human judgement: justice requires a human duty to ‘consider the evidence, deliberate alternative interpretations and reach an informed opinion’. The structure of law and the processes of justice require the presence of a human as a legal agent, and the case against AWS is both a legal and a moral one.

Amoroso and Tamburrini (2017) distinguish two kinds of argument made about AWS: deontological and consequentialist. They identify three main deontological arguments: (a) that AWS would be unable to conform to IHL and IHRL rules governing the use of lethal force; (b) that AWS would create an accountability gap; and (c) that deployment of AWS would be contrary to human dignity and the requirement that ‘the taking of human life should be reserved to human decision-makers’.

In terms of consequentialist arguments, Amoroso and Tamburrini contrast narrow and wide consequentialist reasons. As Tamburrini (2016) points out, narrow consequentialist advantages are sometimes claimed for the future use of AWS—including reductions of casualties due to more accurate targeting and freedom from human self-preservation concerns (Arkin 2009). However, Tamburrini identifies a wide consequentialist view that takes into account the expected effects on peace stability, and on incentives to start wars. He argues that AWS are ‘potentially more threatening to global security than many other conventional weapons’ (Tamburrini 2016, p. 140). He agrees with Sharkey (2008) that their use could reduce the risks of a ‘body bag count’, and as a consequence, remove a major disincentive for war. Sharkey (2012a, b) also highlights concerns about an increase in the pace of war as a result of deploying AWS, and the likelihood of unpredictable interactions between different computational algorithms. Tamburrini (2016) argues that swarms of AWS could weaken traditional nuclear deterrent factors based on mutually assured destruction (swarms of AWS could be used to deliver destructive attacks on strategic nuclear and eliminate an opponent’s second strike nuclear capability, increasing preference for first strike strategies). Amoroso and Tamburrini (2017) point out that AWS even without the ‘lethal’ element, if used to destroy buildings or infrastructure, would still have a global destabilising effect.

On the basis of the preceding account, three main categories of argument against AWS are identified here:

- (i) Arguments based on technology and the current and likely near future abilities of AWS to conform to IHL (i.e. what they *can* do).
- (ii) Deontological arguments based on the need for human judgement and meaningful human control of lethal and legal decisions, and on considerations of what AWS *should* do. These include arguments based on the concept of human dignity.
- (iii) Consequentialist reasons about their effects on the likelihood of going to war. These reasons include political arguments about their effects on global security, and are not necessarily labelled as consequentialist.

In contrast to Amoroso and Tamburrini (2017), arguments based on the abilities of AWS (i), are distinguished here from arguments about the need for human judgement (ii). This is in order to place an increased focus on what weapons systems and artificial intelligence are capable of, and on the differences between humans and computational artifacts or machines. The possibility that AWS will one day be able to conform to IHL is sometimes raised by those in favour of a ban (e.g. Asaro 2012; Heyns 2017). However this requires that they become moral agents, capable of empathic concern for humans and a real understanding of human social behaviour. There are those, including the present author, who question whether this will ever be possible (Sharkey 2017; Hew 2014; Johnson and Miller 2008).

Since the focus of this paper is on a consideration of the particular claim that AWS should be banned because they are against human dignity, we turn now to look at this idea in more detail.

Human dignity and lethal autonomous weapons

A number of writers and campaigners have stated that autonomous lethal weapons are against the notion of human dignity. In a Human Rights Watch report, Goose writes ‘Fully autonomous weapons would also undermine human dignity, because as inanimate machines they could not understand or respect the value of life, yet they would have the power to determine when to take it away’ (Goose 2017). Docherty (2014) also presents arguments against autonomous weapons, and states, ‘fully autonomous weapons could undermine the principle of dignity, which lies at the heart of international human rights law and declares that every human is worthy of respect. An inanimate machine could not truly respect the value of a human life or comprehend the significance of its loss. Allowing a machine to make determinations about when to take life away would vitiate the importance attached to such decisions and degrade human dignity’.

These quotations emphasise the idea that it is an affront to an individual’s dignity if the decision to kill them is made by a machine that does not recognise the value of their life. Heyns (2017) similarly asks, ‘Is it not an affront to human dignity if robots have the power of life and death over humans?’ He offers a number of reasons for this being the case. Amongst them is that the person targeted by AWS is reduced to being ‘an object that has to be destroyed’, where there is no possibility of appealing to the humanity of the enemy. The use of autonomous weapons would remove the potentially restraining influences of humanity. As well as the dignity of those killed or attacked by autonomous weapons, Heyns also claims that the dignity of those in whose name such attacks are carried out is compromised, because the opportunity to be a moral person and to make moral decisions, is lost when machines are used to make lethal decisions.

Heyns’ arguments are based on what he holds to be the human right to a dignified life. Ulgen (2016) also argues that AWS are against human dignity, since they go against many of the central tenets of Kant’s account of human dignity. For one, the use of autonomous weapons denies the equality of persons since combatants using them are removed from physical risk at the same time as their targets are exposed to an increased risk. For another, autonomous weapons ‘diminish the duty not to harm others’. Like Heyns (2017), Ulgen insists that the use of such weapons ‘would devalue humanity by treating humans as disposable inanimate objects rather than ends with intrinsic value and rational thinking capacity’ (Ulgen 2016). They could increase the suffering and humiliation of targets—for example, certain Hellfire missiles cause burning and incineration of bodies. Similarly, the continued threat and use of autonomous weapons can create unacceptable stress and psychological harm in the civilian population.

Johnson and Axinn (2013) also state that ‘To give a programmed machine the ability to ‘decide’ to kill a human is to abandon the concept of human dignity’ since to do so is to treat a rational being as an object. They insist that a machine cannot be moral, but can only follow the values of its programmers.

Asaro (2012) makes the case that to preserve human morality, justice and law, autonomous lethal systems must not be accepted, and concludes, ‘As a matter of the preservation of human morality, *dignity*, justice, and law we cannot accept an automated system making the decision to take a human life’ (Asaro 2012, p. 708, emphasis added). Bhuta et al. (2016) also emphasise the relationship between human dignity and law, and refer to what they term ‘the principle of human dignity’ in their discussions of AWS, pointing out its roots in human rights law, and in IHL. They cite the ‘Martens Clause’ and its modern occurrence in Article 1(12) of Additional Protocol 1, according to which civilians and combatants remain under ‘the principles of international

law derived from established custom, from the principles of humanity and from the dictates of public conscience' (p. 377). They state that AWS would be against human dignity because allowing machines (or objects) to make life or death decisions would compromise the value of each individual as a subject of law.

A report written for the Böll Foundation (Amoroso et al. 2018) distinguishes between two kinds of argument that AWS are against human dignity: arguments centred on agent relevant duties and arguments centred on patient relevant rights. The report describes the agent relevant argument as being based on the need for the act of killing to be grounded in human judgement, since suppressing a human life can only be legally justified if it is based on a considered and informed decision. The patient relevant argument is that human dignity would be denied if people were subject to robotic lethal decision making which gave them no possibility of appealing to the humanity of the enemy. The distinction between the two kinds of argument is interesting, and constitutes a step towards being more explicit about the reasons for claiming that AWS are against human dignity. However, the report is focussed on providing an overall account of the risks of AWS for German foreign and security policy, and as such provides only limited consideration of human dignity, and does not address any of the criticisms of and questions about human dignity that have been raised.

Criticisms of human dignity claims

Some concerns have been raised about the merits of placing a strong emphasis on human dignity in arguments against autonomous weapons (Birnbacher 2016; Lin 2015; Saxton 2016; Pop 2018). Lin (2015) raises various questions about whether AWS, or killer robots, violate the human right to life, and the right to human dignity. He asks whether they should be held to be against the Martens clause and the public conscience. Although he raises concerns, his conclusion is merely that further work is needed to clarify the meaning of dignity, the Martens clause, and human control and autonomy (see Human Rights Watch 2018 for a recent report on the Martens clause and AWS).

Saxton (2016) discusses what he terms the 'problem with the human dignity argument' about killer robots. He claims that arguments based on human dignity often 'fail to grasp the complexity of evaluating human dignity in warfare'. According to him, the use of autonomous weapons should not be viewed as a violation of human dignity 'due solely to the weapon's autonomy', although his objections to what he terms the human dignity argument are not explicitly stated. He acknowledges that autonomous weapons could remove the possibility of empathy and compassion between combatants, and could therefore result in greater killing and suffering as the 'moral distance' between the military and

their targets is increased. His objection seems to be to claims that to be killed by an autonomous weapon rather than by a human is necessarily against human dignity. He points out that this is not the only way that human dignity can be compromised: it is generally compromised in war when humans are sacrificed to achieve military objectives. His argument is that autonomous weapons threaten human dignity by 'potentially changing the dynamic between weapons and their operators'. He is concerned not to lose the potential advantages that automation could bring to warfare, and argues against the need for a ban. Instead he suggests that further thought and investigation is needed to ensure that enough human control of weapons is maintained to ensure that humans can remain responsible and accountable for their use. Although his objections to arguments based on human dignity are not clear, he raises pertinent issues, such as the idea that human dignity is generally compromised by warfare, and not only by autonomous weapons.

Birnbacher (2016) has been the main critic of the emphasis on human dignity in arguments against AWS, and his arguments are more explicit. He acknowledges the many risks and problems of AWS, but disagrees with particular claims about the threats to human dignity. He objects to what he terms 'inflationary' uses of the term. In Birnbacher's account, the term 'human dignity' should be applied only to the individual and not to the human species as a whole. As a consequence, he complains about examples in which 'dignity' is applied to the human species, as it is when it is claimed that delegating the kill decision to machines is against human dignity. This complaint is a little contentious, since as will be discussed in the next section, there is interest in the idea of collective dignity. It could also be argued that each individual delegation of a kill decision to a machine is against that individual's human dignity, and that there is no need to assume that in such instances 'dignity' is being applied to the human species as a whole.¹

Birnbacher also objects to a tendency to equate human dignity with the whole of morality. This, he argues, is a problem because it weakens the concept of dignity and would mean that every immoral act was against human dignity. A further objection that he raises is to the use of the term as the expression of an emotional reaction and as a rhetorical device without any further specific meaning.

For Birnbacher, human dignity only applies to the individual and implies a set of basic human rights. His 'tentative' list (Birnbacher 2016) consists of the following: (1) the right not to be severely humiliated and made the object of public contempt; (2) the right to a minimum of freedom of action and decision; (3) the right to receive support in situations of severe need; (4) the right to a minimum quality

¹ A point made by an anonymous reviewer of this paper.

of life and relief of suffering; (5) the right not to be treated merely as a means to other people's ends, i.e. without consent and with severe harm or risk of harm. Based on this reasoning, he considers the question of whose individual dignity might be affected by the deployment of autonomous weapons and how it might be affected. He concludes that the most likely candidate is the dignity of individual civilians (not the dignity of combatants since soldiers already know what is involved in war and may have the possibility of opting out). Civilians cannot usually remove themselves from the situation, and are at risk from direct attack, or incidental losses from attacks aimed at military targets. However, when Birnbacher considers *how* autonomous weapons could directly affect the dignity of civilians, he disagrees with Heyns' statement that 'giving machines greater power to take life and death decisions is demeaning' and that AWS should therefore be taken to violate the 'right to dignity' (Heyns 2017). Birnbacher does not accept that the use of lethal autonomous weapons per se directly results in civilians being demeaned or humiliated, except through the creation of mental pain and subjective suffering. He does agree that AWS pose threats to civilians by preventing even a minimal quality of life and that to deploy such weapons against them is to treat them as a means to an end that does not reflect their own aims and interests.

Birnbacher points out a number of features of autonomous weapons that are likely to cause severe dread in civilians. These include the usually asymmetrical nature of their use, where only one side in the conflict possesses such weapons. Then there is the unpredictability and inscrutability of the weapons, and their limited capacity to discriminate between combatants and non-combatants and to observe the rules of proportionality. All these are reasons why the threat of an attack by AWS could create such intense 'mental pain' as to be contrary to the human dignity of their actual and potential victims. Such mental pain could be held to be against human dignity in the same way that torture is. However such pain, fear and anxiety are not unique features of AWS, and could also apply to other weapons, such as remote controlled missiles and drones, depending on their use.

In an ICRC blog, Pop (2018) also considers the relationship between human dignity and anti-AWS arguments. Her complaint is that there is too little reflection on how human dignity is used in such arguments to warrant the conclusion that allowing a machine to decide to kill a human is against human dignity. This is an important point, and one that coheres with the argument of this paper, although a wider selection of accounts of dignity is considered here. Pop identifies two interpretations of dignity and considers their relationship to arguments against AWS. The first is the interpretation of dignity based on the Kantian notion that human dignity stands for unconditional intrinsic value, which has its source in human autonomy. According to this view, AWS

would affect the human ability to make self-determined choices, and would therefore be against this sense of human dignity. But so too would any force which harms or interferes with human agency. She also identifies a different view of human dignity as noble rank or status, as articulated by Waldron (2009). For Waldron, dignity has its origins in the notion of noble rank and status, but 'the modern notion of *human* dignity involves an upwards equalization of rank, so that we now try to accord to every human being something of the dignity, rank, and expectation of respect that was formerly accorded to nobility' (Waldron 2009, p. 229). Pop suggests that an argument could be made to the effect that to be killed by a machine, a lesser entity, would be against this conception of human dignity. However, as she points out, this is not the argument that is made in anti-AWS arguments. Her main conclusion is that the reasons for claiming that AWS are against human dignity are not sufficiently spelled out. In the absence of a consensus about the meaning of human dignity, she suggests the concept should not be used in the AWS debate.

Differing accounts of human dignity

As well as the criticisms that have been made of the use of human dignity in arguments against AWS, another set of problems stems from a lack of clarity about what human dignity actually is. This lack of clarity has been complained about by a number of authors in a variety of contexts. Several writers have complained that the term is used too frequently. For instance, Macklin (2003) states that it is so frequently invoked in a medical context that its use could be eliminated 'without any loss of content'. Pinker, in a criticism of a report on human dignity and bioethics, derides it as 'a squishy, subjective notion, hardly up to the heavyweight moral demands assigned to it' (Pinker 2008). Werner (2014) describes dignity as 'a heavily contested, multifaceted, and ambiguous concept'. Shultziner (2007) reviews its uses and points out that dignity is used in different ways in many different contexts.

It is recognised that there are contradictions amongst the different uses of the term 'dignity', when it is both talked about as something that can be lost or reduced through humiliating treatment, yet also cited as the basis for the right to be treated humanely. Nordenfelt (2003) points out the paradox of the classic concentration camp example, in which prisoners are seen as being degraded and robbed of their dignity by inhuman treatment, even though it is recognised that everyone is of equal value and has a dignity that cannot be taken away. This paradox can be resolved by distinguishing between the inviolable or universal dignity that is an inherent property of human beings, and other forms of dignity that can be held to varying degrees (Bostrom 2008; Schroeder 2010; Nordenfelt

2004). Inviolable dignity is closely related to the concept of human rights, and cannot be removed by dint of humiliating treatment. Many international documents refer to human dignity as the justification for human rights. For instance, the 1948 Universal Declaration of Human Rights preamble states that ‘the recognition of the inherent dignity and of the equal and inalienable rights of all members of the human family is the foundation of freedom, justice and peace in the world’. The Basic Law of Germany, Article 1, sentence 1, states, “Human dignity is inviolable. To respect and protect it is the duty of all state authority” (1949).

Apart from inviolable dignity, there are other forms of dignity that are attributable to individuals to varying degrees. For Bostrom (2008) these include a form of dignity he terms ‘dignity of status’, which refers to the position and standing of an individual: (Schroeder 2010 uses the term ‘aristocratic dignity’, and; Nordenfelt 2004, ‘dignity of merit’ for what appear to be the same idea). Then there is ‘dignity of quality’ (Bostrom 2008) ‘meritorious dignity’ (Schroeder 2010) and ‘dignity of moral or existential character’ (Nordenfelt 2004), all of which are used to refer to exceptional individuals who act in an honourable way, especially when in adverse circumstances (Nelson Mandela being a paradigmatic example).

Jacobson (2009), in the context of healthcare and human rights, provides a taxonomy of dignity. She reports a qualitative grounded theory exploration of dignity based on semi-structured interviews with persons marginalised by health and social status, or providing services to those persons, and those working in the field of human rights and health. On the basis of previous literature, she distinguishes between inviolable human dignity, and social dignity. She then divides social dignity into dignity of self, and dignity in relation. Dignity of self is a quality of self-respect and self-worth. Dignity in relation ‘refers to the ways in which respect and worth are conveyed through individual and collective behaviour’ (Jacobson 2009). On the basis of interviews, she identifies the ways in which her participants felt their dignity had been violated, or promoted. Individual dignity could be harmed or benefited by injuries or benefits to the self, by violations or respect to the body, and by injuries or benefits to moral agency or personhood. An example of an injury to the self is where a person is treated with contempt—as if they had no value, or excluded from physical or social settings. Collective dignity could be harmed by processes such as discrimination and exclusion, or promoted through recognition and acceptance. Interestingly, Jacobson also identifies some of the long-term consequences of dignity violations whereby an individual experiences a series of losses, including loss of self-worth, loss of moral standing, and loss of confidence. The longer-term effects on individuals include social marginalisation, passivity and chronically poor physical and

mental health. The collective effects include group traumatisation, and a loss of dignity.

Not only is it possible to identify different forms of dignity, there are also cultural variations in the way in which the concept of dignity can be interpreted. Shultziner (2007) points out how the set of values implied by the idea of human dignity can differ between societies. For instance, an Islamic interpretation of human dignity may be quite different to that of a Western liberal-democracy—involving for instance a different view of materialism, and of the rights and social position of men and women. The values implied by human dignity have also changed over time. Thus from the nineteenth century, slavery has been seen as an affront to human dignity, but the institution of slavery has a long history starting with the ancient civilisations of Mesopotamia, India, China and Greece and was not always viewed as we view it now. There can be different views of dignity within contemporary society also: for instance, many see rights to abortion and euthanasia as representative of human worth and dignity, whilst others see them as an affront to human dignity.

In addition to cultural and historical differences, the meanings attributed to dignity vary according to context. The meanings of dignity in a health context are likely to be quite different to the meanings of dignity in a war zone. In health care, as discussed by Sharkey (2014), concerns about dignity are often related to bodily functions, access to toileting facilities, and being addressed respectfully. In a war zone, the interpretation of dignity is likely to be quite different, and overriding concerns about death, suffering, and stress will leave little room for worries about personal hygiene.

Questions have also been raised about the extent to which there is such a thing as collective dignity, or whether dignity is only something that can apply to the individual, as is argued by Birnbacher (2016). Werner (2014) points out a growing number of references to collective dignity (e.g. De Gaay Fortman 2011; Falk 2009), but claims that the term is often used without any attempt to clarify its meaning. One of the uses of ‘collective dignity’ that Werner points out refers to the duties of a collective, for instance the duties of a government or similar organisation. Another use is when human rights are being emphasised: for instance where the collective dignity of an identifiable group is said to have been reduced when one member of the group is humiliated. A third use is the idea of collective human dignity that is shared among the members of the human species. This form of collective human dignity is more often referred to in the context of bioethics, where it might be claimed for instance, that human cloning was against the dignity of the species. Werner agrees that it is conceptually possible to ascribe dignity to a collective, in the sense that the collective has certain rights or duties. However he argues against

the usefulness of this. He claims that although membership of a collective might invoke certain rights or duties, it is not clear that these can be usefully described as collective rights or duties—or whether they should be seen as individual.

Although it is clear that human rights and dignity are closely related, both historically and conceptually, there is also some confusion and disagreement about whether dignity is itself a right, or whether it forms a justification for human rights. Heyns (2017) talks of the right to human dignity, but in a review of its uses, Shultziner (2007) claims that dignity is more often used as a ‘bedrock truth justification’ for human rights than as a right in itself. For Birnbacher (2016), as stated earlier, human dignity should only be applied to the individual, and not to the human species as a whole. As described earlier, he proposes that the human dignity inherent in the individual implies a set of basic human rights which include; the right to a minimum of freedom of action and decision; the right to a minimum of quality of life and relief of suffering; and the right not to be treated merely as a means to other people’s ends. He adds that privacy should also be considered a right. He emphasises that only a minimum of each right is covered by human dignity. In this, his set of basic human rights bears some similarity to a related list provided by Nussbaum (2006, 2011) in her account of the capability approach (CA).

Nussbaum (2006, 2011) uses the concept of dignity as the basis for a list of human rights or capabilities. Her concept of human dignity takes a wider perspective. She argues that for a life worthy of human dignity, all humans should be able to achieve at least a threshold level of ten central capabilities. The full list can be found elsewhere (Nussbaum 2006, pp. 24–33), but it includes items such as (1) life: being able to live to the end of a life of normal length; not dying prematurely, or before one’s life is so reduced as to be not worth living; (3) bodily integrity. Being able to move freely from place to place; to be secure against violent assault, including sexual assault and domestic violence; having opportunities for sexual satisfaction and for choice in matters of reproduction; (5) emotions. being able to have attachments to things and people outside ourselves; to love those who love and care for us, to grieve at their absence; in general, to love, to grieve, to experience longing, gratitude, and justified anger. Not having one’s emotional development blighted by fear and anxiety, and (10) control over one’s environment. The list of ten capabilities is described by Nussbaum (2006) as ‘open-ended and subject to on-going rethinking and revision, in the way that any society’s account of its most fundamental entitlements is always subject to supplementation (or deletion)’ (p. 78). The CA has been found to provide a useful framework by a number of writers on robot ethics (Coeckelbergh 2010; Borenstein and Pearson 2010; Sharkey 2014; Vallor 2011), but its relevance to military action, and to AWS does not yet seem to have been explicitly discussed.

Nonetheless, it provides a comprehensive account of human dignity, and as such is relevant here.

In summary, it should be apparent that not only have some specific questions been raised about the impact of AWS on human dignity, but also that there is a lack of a clear consensus about what dignity is. There are questions about whether collective dignity is a meaningful concept, or whether dignity is something that applies only to the individual. Different contexts lead to differing emphases on aspects of dignity, and views of dignity vary with culture and over time. Sometimes dignity is assumed to be a human right itself (Heyns 2017), and sometimes it is seen as the basis for human rights (Birnbacher 2016; Nussbaum 2006).

Questions about human dignity and AWS

Now that we have looked at the main arguments against AWS, the claims that made by some that they compromise human dignity, and some problems that have been raised about these claims, it is time for further reflections. These reflections will be organised around a set of questions emerging from the previous discussions:

1. What are the main ways in which AWS can be said to threaten human dignity?
2. Are there advantages for the campaign against killer robots in claiming that AWS are against human dignity?
3. Are AWS against human dignity in a way that other weapons are not?
4. Are claims based on human dignity the best way to argue against AWS?

Q1: What are the main ways in which AWS can be said to threaten human dignity?

An account of the claims made about AWS and human dignity was provided in “[Human dignity and lethal autonomous weapons](#)” section. They are summarised and extended here. It is apparent from the preceding discussions that there are a number of different ways in which AWS can be said to be against human dignity. A particular claim about AWS and human dignity has been described in a UNIDIR report as ‘being at the core of the concerns raised about fully autonomous weapons’ (UNIDIR 2015). The claim is that allowing a weapon system to make a kill decision is against human dignity because weapons, computers and robots are unable to understand or respect the value of life, or understand the significance of its loss (Docherty 2014; Goose 2017; Heyns 2017).

AWS have also been held to be against human dignity because they cannot conform to the laws of war, and because they replace the human reflection that is essential

for justice, morality and law. Human interpretation of law is both assumed and required. This argument is made by Asaro (2012), and by Heyns (2017) when he claims that a lack of human deliberation would render any lethal decisions arbitrary and unaccountable.

There are also a number of ways in which AWS can be shown to be against Kantian dignity, as shown by Ulgen (2016). They can reduce the equality of persons, especially since they are often used only one side in a conflict. They diminish the duty not to harm others. They can also increase suffering and humiliation and create unacceptable psychological stress.

Birnbacher (2016) takes the view that human dignity implies a set of basic human rights, which include the right to a minimum of freedom of action and decision, the right to a minimum level of quality of life and relief of suffering and the right not to be treated as a means to other people's ends. AWS could affect all of these: limiting freedom, reducing quality of life, and creating suffering.

Birnbacher's view is related to Nussbaum's version of the CA, in which it is held that a life worthy of human dignity requires a threshold level of a set of 10 central capabilities. If we extrapolate from the CA, since its relevance to AWS has not yet been established, it is apparent that AWS would have a negative impact on several of the central capabilities, including (1) living a life of normal length, (3) being able to move freely from place to place, (5) not having one's emotional development blighted by fear and anxiety, and (10) control over one's environment. Thus within the context of the CA, AWS can be seen to be against human dignity because they would have a negative effect on people's access to at least four central capabilities that are essential for a life worthy of human dignity.

It seems that there are several ways in which AWS can be seen to reduce or diminish human dignity. Are there any ways in which they could be claimed to enhance it? It is difficult to see how this could be the case. In a different context, Jacobson (2009) reported some examples of processes that seem to promote dignity in persons marginalised by health and social status; for example, empowerment (working with others to enhance their capacities, capabilities and competencies); recognition (acknowledging the humanity of others by paying attention and showing appreciation); and Levelling (minimising asymmetry). However AWS seem unlikely candidates for promoting any of these, and indeed seem to do the opposite. Lin (2015) hints at one way in which AWS might result in an increase in human dignity compared to other weapons when he mentions the possibility that autonomous weapons might not need to be lethal at all since they do not need to protect their own lives, and could wound combatants rather than killing them. It might also be argued that Artificial Intelligence, robots and their sensors could be used to improve awareness of what is happening in

the 'fog of war' and to reduce civilian casualties, but such information could be made available to human commanders, and does not itself legitimise the use of AWS.

The fact that it is possible to identify a variety of ways in which AWS can be held to be against human dignity illustrates and reinforces reasons for supporting a ban. It also indicates a problem with using dignity as the basis for arguments against them. The underlying difficulty is the lack of consensus about the meaning of dignity—although the preservation of human dignity is generally recognised as essential, what this means varies between cultures, contexts, historical era, and philosophical position.

Q2: Are there advantages for the campaign against killer robots to claiming that AWS are against human dignity?

There could be some campaigning advantages. Saying that something is against human dignity evokes a strong visceral response. Even though dignity is difficult to define clearly, people have an intuitive understanding of its meaning, and of the importance of maintaining and preserving it. Reference to human dignity can highlight a repugnance to the idea of machines having the power of life or death decisions over humans, as highlighted in a UNIDIR report (2015).

Claiming that AWS are against human dignity accentuates and underlines concerns about these weapons. A statement by the university president of the South Korean university, in response to a letter from an international group of scientists calling for a boycott of the university, illustrates this: 'I reaffirm once again that KAIST will not conduct any research activities *counter to human dignity* including autonomous weapons lacking meaningful human control'. (emphasis added: The Guardian, "'Killer robots': AI experts call for boycott over lab at South Korea University" 5th April 2015).

Q3: Are AWS against human dignity in a way that other weapons are not?

Answering this question is important in order to establish whether it can be argued that AWS, that are not subject to meaningful human control, are against dignity when other weapons that cause death and suffering are not. However, it is not clear how this question should be answered—and this illustrates again a problem with basing arguments for a ban exclusively on human dignity. It is possible to identify at least three answers to this question.

First it can be claimed that war and killing in general are against human dignity. Hasenclever (2014) describes war as a moral evil and states that 'human dignity requires the abrogation of war'. For those adopting this position, the deployment of AWS in battle would be contrary to human dignity,

but so too would any lethal weapon or weapons system, even those entirely operated and controlled by humans.

A second answer would be to acknowledge the inevitability of weapons and war, but to contend that any weapons that caused extreme suffering and humiliation were against human dignity. This would mean that AWS could be argued to be against human dignity, but so too would other weapons such as chemical weapons, and nuclear attack. Birnbacher (2016) holds this position, and argues that AWS and other weapons such as remote controlled missiles could have a negative impact on the dignity of individuals: interfering with their basic human rights by causing subjective pain and mental suffering. Similarly, it is argued here in the context of the CA that weapons that have a negative impact on access to central capabilities should be seen as affecting the ability to lead a life worthy of human dignity. Under a Kantian view of dignity, it can also be argued that as well as AWS, any weapons that cause extreme suffering should be considered to be against human dignity (Ulgen 2016).

A third response is to agree that AWS are against human dignity in a way that other weapons are not. As we have seen, it has been suggested that ‘death by algorithm’ crosses a moral line and is against human dignity because AWS will be unable to understand or value the human lives that they were taking (Goose 2017; Docherty 2014; Heyns 2017).

Q4: How important are claims based on human dignity for arguments against AWS?

This question is central to our consideration of the relationships between dignity and AWS. In the light of the issues reviewed in this paper, it is concluded here that although dignity is an important concern when considering AWS, it is not sufficient on its own. It works best as part of a cluster of other arguments about legal compliance, global security and technical competence. The main problem is, as we have seen, the lack of consensus about the meaning of human dignity. There are questions about whether it makes sense to talk of affronts to collective dignity (Werner 2014), or whether the term dignity should only be applied to individuals. Views of dignity can also be shown to change with time and context (Shultziner 2007). There are also differing views as to whether dignity should be viewed as a human right (Heyns 2017), or as the basis for human rights (Birnbacher 2016).

The lack of clarity associated with claims about AWS and human dignity can be further illustrated by reconsidering the statement from the KAIST president, quoted earlier, in which he affirms that KAIST will not conduct any research activities ‘counter to human dignity’ (The Guardian, “‘Killer robots’: AI experts call for boycott over lab at South Korea University” 5th April 2015). It is clear that the statement assumes that AWS are against human dignity—what is not

clear is what other research activities will be avoided in the future. A strong interpretation could be that the university will not undertake any further military research. An even stronger one would be that KAIST will avoid research in which artificial intelligence is used to replace meaningful human control. Or perhaps the statement means that the university will not research weapons that do not conform to IHL. Although the statement has a pleasing resonance to it, its meaning is unclear because of the ambiguity associated with the concept of human dignity.

Concluding that arguments based on human dignity are not the best way to argue against AWS is not at all the same as saying that such weapons are acceptable. As explained earlier, there are strong reasons to oppose their use and to argue for a ban. A complete review of all of the reasons and arguments against AWS is beyond the scope of this paper, but three main categories of argument were identified: (i) arguments based on technology and the current and likely near future abilities of AWS to conform to IHL; (ii) arguments based on the need for human judgement and meaningful human control of lethal decisions; and (iii) arguments about the expected effects of AWS on the likelihood of going to war and on global instability.

Of course, deciding which of these categories of argument are the most convincing will depend on who is being addressed. Politicians might well be more convinced by arguments in the third category based on the effects of AWS on global instability. Philosophers and lawyers would probably be more convinced by arguments in the second category about the need for human judgement in matters of law and justice. Artificial Intelligence researchers might be more interested in the first category. There is also the possibility of combining the arguments, instead of selecting the most appropriate. Amoroso and Tamburrini (2017) present an interesting idea for combining deontological and consequential reasons by means of a confluence model that resolves potential conflicts between deontological and consequential arguments by prioritising deontological arguments.

For this author, the most convincing arguments for a ban of AWS are those in the first category that focus on the current and foreseeable future abilities of robots and computer systems. AWS are unable to reliably discriminate between civilians and combatants, both because it is not possible to formulate a set of rules that will always enable such discrimination, but also because they do not have the necessary situational awareness and understanding of human actions and intentions. This lack of situational awareness and understanding also means that the principles of proportionality and military understanding are beyond them. These require informed and considered human judgement, and a moral understanding and competence. As discussed by Sharkey (2017), not only do current machines not have these abilities, there is also no good reason to expect that machines can be

either programmed or trained to develop an understanding of what is and is not morally acceptable. As well as their relevance for arguments in favour of a ban on AWS, the limitations of computational artifacts, and the differences between living and artificial machines, also imply the need to limit or prohibit the use of robots in other spheres of life where their use would have a significant impact on humans.

Conclusions

In this paper, we have examined the relationship between human dignity and AWS, or ‘killer robots’. We have looked at the main arguments that have been made against the use of these weapons, and distinguished between three categories of argument. We have considered the claims that have been made that the use of AWS is against human dignity, the objections made to these claims, and the lack of consensus about the meaning of ‘human dignity’.

A set of four questions were then identified and addressed in an effort to clarify the relationship between AWS and human dignity.

- Q1. What are the main ways in which AWS can be said to threaten human dignity?
- Q2. Are there advantages for the campaign against killer robots in claiming that AWS are against human dignity?
- Q3. Are AWS against human dignity in a way that other weapons are not?
- Q4. How important are claims based on human dignity for arguments in favour of a ban on AWS?

On the basis of the foregoing reviews and discussions, it is clear that AWS should be considered to be against human dignity. In answer to Q1 [What are the main ways in which AWS can be said to threaten human dignity?], several reasons for claiming that AWS are against human dignity were identified. Q2 [Are there advantages for the campaign against killer robots in claiming that AWS are against human dignity?] was answered by agreeing that there could be some campaigning advantages to claiming that AWS are an affront to human dignity: even though it may not be clear what is meant by saying that something is against human dignity, it can still function as a rallying cry.

The response to Q3 [Are AWS against human dignity in a way that other weapons are not?] was to point out that even though it can be readily concluded that AWS are against human dignity, they are not unique in this respect. There are other weapons that cause pain and suffering which should also be viewed as being against human dignity. There are also other uses of robots which could create

suffering, or limit individuals’ access to aspects of life necessary for a life worthy of human dignity. For instance, Sharkey (2014) argues that there are uses of robots, such as employing them for the exclusive care of older people, which should be considered to be against the view of human dignity represented by the CA.

There are also other uses of technology that can be seen as being against human dignity. Heyns (2017) looks beyond warfare and raises concerns about where we allow machines to make decisions. He states that allowing machines to make non-lethal decisions that affect humans is also against human dignity, writing, “The notion of ‘meaningful human control’ should be developed as a guiding principle not only for the use of autonomous weapons, but for the use of artificial intelligence in general; not merely focussing on isolated uses of such technologies but on the role of technology as such in our future. Allowing technology not only to supplement but indeed to replace human decision-making will undermine the very reason why life is valuable in the first place.” (Heyns 2017). His argument is that we should be concerned about replacing human decision-making with machines, and asks how far the process of transferring power from humans to machines should go. He suggests that autonomous weapons, with life or death stakes, are a pivotal test case.

It seems then that there are many technological applications that can be considered to impact the human dignity of individuals. It is also the case that human behaviour can have a negative effect on individual dignity. However if it is accepted that there are many weapons, artifacts, and human behaviours that are held to be against human dignity, then this itself becomes a reason for not relying too heavily on human dignity in arguments against AWS, as distinct from other means and weapons of warfare. Another reason stems from the previously identified lack of consensus about, and the differing accounts of what human dignity actually is.

So after these reflections, the final conclusions drawn here are that human dignity can indeed be said to be compromised by the use of AWS, but also that there are many different interpretations of human dignity and many different ways in which it can be affected. The answer given to Q4 [How important are claims based on human dignity for arguments against AWS?] is essentially that the risk to human dignity is only one of many reasons for calling for a ban of autonomous weapons and for insisting on the need for meaningful human control of lethal weapons in war; and it is not the most compelling. Three categories of reasons for opposing AWS were identified here. In opposing the development and use of AWS it makes sense to be able to draw on them all, combining them or choosing the most relevant, and not relying exclusively on any one of them.

Acknowledgements I would like to thank Bonnie Docherty and Noel Sharkey for their comments on an earlier version of this paper.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Amoroso, D., Sauer, F., Sharkey, N., Suchman, L., & Tamburrini, G. (2018). Autonomy in weapon systems: The military application of artificial intelligence as a litmus test for Germany's new foreign and security policy. In *Heinrich Böll Stiftung publication series on democracy* (Vol. 49).
- Amoroso, D., & Tamburrini, G. (2017). The ethical and legal case against autonomy in weapons systems. *Global Jurist*. <https://doi.org/10.1515/gj-2017-0012>
- Arkin, R. C. (2009). *Governing lethal behavior in autonomous robots*. Boca Raton: CRC Press.
- Asaro, P. (2012). On banning autonomous lethal systems: Human rights, automation and the dehumanizing of lethal decision-making, special issue on new technologies and warfare. *International Review of the Red Cross*, 94(886), 687–709.
- Bhuta, N., Beck, S., & Geiss, R. (2016). Present futures: Concluding reflections and open questions on autonomous weapons systems. In N. Bhuta, S. Beck, R. Geiß, H. Liu, & C. Kreß (Eds.), *Autonomous weapons systems: Law, ethics, policy* (Vol. 15, pp. 347–383). Cambridge: Cambridge University Press
- Birnbacher, D. (2016). Are autonomous weapon systems a threat to human dignity? In N. Bhuta, S. Beck, R. Geiß, H. Liu, & C. Kreß (Eds.), *Autonomous weapons systems: Law, ethics, policy* (pp. 105–121). Cambridge: Cambridge University Press.
- Borenstein, J., & Pearson, Y. (2010). Robot caregivers: Harbingers of expanded freedom for all? *Ethics and Information Technology*, 12(3), 277–288.
- Bostrom, N. (2008). Dignity and enhancement. In *Human dignity and bioethics: Essays commissioned by the council on bioethics* (pp. 173–207). Washington, DC: The President's Council on Bioethics.
- Coeckelbergh, M. (2010). Health care, capabilities, and AI assistive technologies. *Ethical Theory and Moral Practice*, 13(2), 181–190.
- De Gaay Fortman, B. (2011). Minority rights: A major misconception? *Human Rights Quarterly*, 33(2), 265–303.
- Docherty, B.L. (2014). *Shaking the foundations: The human rights implications of killer robots*. <https://www.hrw.org/report/2014/05/12/shaking-foundations/human-rights-implications-killer-robots>.
- Falk, R. A. (2009). *Achieving human rights*. New York: Routledge.
- Goose, S. (2017). The growing international movement against killer robots. *Harvard International Review*. Accessed May 17, 2018 from <http://hir.harvard.edu/article/?a=14022>.
- Hasenclever, A. (2014). Human dignity and war. In M. Düwell, J. Braarvig, R. Brownsword, & D. Mieth (Eds.), *The Cambridge handbook of human dignity: Interdisciplinary perspectives* (pp. 439–445). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511979033.054>.
- Hew, P. C. (2014). Artificial moral agents are infeasible with foreseeable technologies. *Ethics and Information Technology*, 16, 197–206.
- Heyns, C. (2013). *Report of the Special Rapporteur on extrajudicial, summary or arbitrary executions, A/HRC/23/47*. New York: United Nations.
- Heyns, C. (2016). Autonomous weapons systems: Living a dignified life and dying a dignified death. In N. Bhuta, S. Beck, R. Geiß, H. Liu, & C. Kreß (Eds.), *Autonomous weapons systems: Law, ethics, policy* (pp. 3–20). Cambridge: Cambridge University Press
- Heyns, C. (2017). Autonomous weapons in armed conflict and the right to a dignified life: An African perspective. *South African Journal on Human Rights*, 33(1), 46–71.
- Human Rights Watch. (2012). *Losing humanity: The case against killer robots*. Accessed May 17, 2018, from <http://www.hrw.org/reports/2012/11/19/losing-humanity-0>.
- Human Rights Watch. (2014). *Shaking the foundations: The human rights implications of killer robots*. <https://www.hrw.org/report/2014/05/12/shaking-foundations/human-rights-implications-killer-robots>.
- Human Rights Watch. (2018). *Heed the call: A moral and legal imperative to ban killer robots*. <https://www.hrw.org/report/2018/08/21/heed-call/moral-and-legal-imperative-ban-killer-robots>.
- ICRC (2014). ICRC, autonomous weapon systems: Technical, military, legal and humanitarian aspects. In *Expert meeting*. Geneva, Switzerland, 26–28 March 2014, 1 November 2014, 3.
- Jacobson, N. (2009). A taxonomy of dignity: A grounded theory study. *BMC International Health and Human Rights*. <https://doi.org/10.1186/1472-698X-9-3>
- Johnson, A. M., & Axinn, S. (2013). The morality of autonomous robots. *Journal of Military Ethics*, 12(2), 129–141.
- Johnson, D. G., & Miller, K. W. (2008). Un-making artificial moral agents. *Ethics and Information Technology*, 10, 123–133.
- Lin, P. (2015). Do killer robots violate human rights? *The Atlantic*. Accessed May 17, 2018 from <https://www.theatlantic.com/technology/archive/2015/04/do-killer-robots-violate-human-right/s/390033/>.
- Macklin, R. (2003). Dignity is a useless concept. *British Medical Journal*, 327, 1419–1420.
- Nordenfelt, L. (2003). Dignity of the elderly: An introduction. *Medicine, Health Care and Philosophy*, 6(2), 99–101.
- Nordenfelt, L. (2004). The varieties of dignity. *Health Care Analysis*, 12(2), 69–81.
- Nussbaum, M. C. (2006). *Frontiers of justice*. Cambridge: Belknap Press.
- Nussbaum, M. C. (2011). *Creating capabilities: The human development approach*. Belknap Press: Harvard University.
- Pinker, S. (2008). *The stupidity of dignity*. Washington, DC: The New Republic.
- Pop, A. (2018). *Autonomous weapons systems: A threat to human dignity?* Humanitarian Law and Policy, ICRC blog. Consulted May 17, 2018 from <http://blogs.icrc.org/law-and-policy/>.
- Saxton, A. (2016). *(Un)dignified killer robots? The problem with the human dignity argument*. Lawfare blog. Consulted May 17, 2018 from <https://www.lawfareblog.com/undignified-killer-robots-problem-human-dignity-argument>.
- Schroeder, D. (2010). Dignity: One, two, three, four, five, still counting. *Cambridge Quarterly of Healthcare Ethics*, 19(1), 118–125.
- Sharkey, A. (2014). Robots and human dignity: A consideration of the effects of robot care on the dignity of older people. *Ethics and Information Technology*, 16, 63–75. <https://doi.org/10.1007/s10676-014-9338-5>. 1.
- Sharkey, A. (2017). Can we program or train robots to be good? *Ethics and Information Technology*. <https://doi.org/10.1007/s10676-017-9425-5>.
- Sharkey, N. (2008). Grounds for discrimination: Autonomous robot. *RUSI Defence Systems*, 11, 86–89.
- Sharkey, N. (2012a). The evitability of autonomous robot warfare. *International Review of the Red Cross*, 94(886), 787–799.

- Sharkey, N. (2012b). Automating warfare: Lessons learned from the drones. *Journal of Law, Information and Science*, 21(2), 140.
- Shultziner, D. (2007). Human dignity: Functions and meanings. In J. Malpas & N. Lickiss (Eds.), *Perspectives on human dignity: A conversation*. Dordrecht: Springer.
- Suchman, L. (2016). Situational awareness and adherence to the principle of distinction as a necessary condition for lawful autonomy. In *Panel presentation at CCW Informal meeting of experts on lethal autonomous weapons*, Geneva, April 12, 2016.
- Tamburrini, G. (2016). On banning autonomous weapons systems: From deontological to wide consequentialist reasons. In N. Bhuta et al. (Eds.), *Autonomous weapons systems: Law, ethics, policy* (pp. 121–141). Cambridge: Cambridge University Press.
- Ulgen, O. (2016). Human dignity in an age of autonomous weapons: Are we in danger of losing an 'elementary consideration of humanity'? 8(9). In *ESIL conference paper series*, 1–19. *European Society of International Law (ESIL) 2016 annual conference (Riga)*. Published on January 31, 2017, ESIL SSRN.
- UNIDIR. (2015). *The weaponization of increasingly autonomous technologies: Considering ethics and social values*. UNIDIR Responses 3.
- Vallor, S. (2011). Carebots and caregivers: Sustaining the ethical ideal of care in the 21st century. *Philosophy & Technology*, 24(3), 251–268.
- Waldron, J. (2009). *Dignity, rank and rights*. Tanner Lectures, UC Berkeley.
- Werner, M. (2014). Individual and collective dignity. In M. Düwell, J. Braarvig, R. Brownsword & D. Mieth (Eds.), *The Cambridge handbook of human dignity: Interdisciplinary perspectives* (pp. 343–352). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511979033.040>.