

(forthcoming article in *Baltic Yearbook of International Law*, vols. 17/18, 2018)

Human Dignity in an Age of Autonomous Weapons: Are We in Danger of Losing an ‘Elementary Consideration of Humanity’?

Ozlem Ulgen¹

Abstract

Military investment in robotics technology is leading to development and use of autonomous weapons, which are machines with varying degrees of autonomy in target, attack, and infliction of lethal harm (that is, injury, suffering or death). Examples of autonomous weapons include weapons systems involving levels of automation and remotely controlled human input, unmanned armed aerial vehicles (UAV), remotely-controlled robotic soldiers, bio-augmentation, and 3D printed weapons. Autonomous weapons generally fall into one of two categories: semi-autonomous, involving some degree of autonomy in certain critical functions such as acquiring, tracking, selecting, and attacking targets, along with a degree of human input or remote control (for example UAV or ‘drones’); and autonomous, involving higher levels of independent thinking as regards critical functions without the need for human input or control (for example, US Navy X-47B UAV with autonomous take-off, landing, and aerial refuelling capability). The trend is clearly towards developing autonomous weapons. Development of new weapons aimed at reducing costs and casualties is not a new phenomenon in warfare. Technological advances have created greater distance between the soldier and the battlefield. A bullet fired from a rifle handled by a human has been superseded by a missile fired from a remotely controlled or autonomous machine. So what makes autonomous weapons different? What particular challenge do they pose international law? Although autonomous

¹ Dr Ozlem Ulgen, Reader in International Law and Ethics, School of Law, Birmingham City University, UK. This was originally presented as a paper at Agora 1 New Forms of Warfare and Armed Conflict of the 12th Annual Conference of the European Society of International Law in Riga, Latvia, on 8 September 2016. I am most grateful to Claus Kress, Ineta Ziemele, and George Ulrich for their insightful comments on an earlier draft.

weapons may be employed to attack non-human targets, such as state infrastructure, here I am primarily concerned with their use for lethal attacks against humans.

In this chapter I focus on autonomous weapons (both semi-autonomous and fully autonomous) and their impact on human dignity under two of Kant's conceptual strands: 1) human dignity as a status entailing rights and duties; and 2) human dignity as respectful treatment. Under the first strand I explore how use of autonomous weapons denies the right of equality of persons and diminishes the duty not to harm others. In the second strand I consider how replacing human combatants with autonomous weapons debases human life and does not provide respectful treatment. Reference is made to contemporary development of Kant's conceptual strands in ICJ and other international jurisprudence recognising human dignity as part of 'elementary considerations of humanity' in war and peace.

1. Human Dignity as an Elementary Consideration of Humanity

New military weapons technology such as autonomous weapons challenge our understanding of what is permissible and impermissible in warfare. Autonomous weapons are characterised by varying degrees of autonomy in the critical functions of acquiring, tracking, selecting, and attacking targets; and, to some extent or even fully, removal of human involvement from the decision-making process to use lethal force. Apart from whether these weapons comply with specific international humanitarian law principles (for example distinction, proportionality, and unnecessary suffering), their characteristics may conflict with a fundamental ethical tenet of humanity that permeates the law, namely, human dignity.

Critics of human dignity refer to it as ‘a deceptive facade’ and ‘an empty space’ open to exploitation and potentially threatening to personal autonomy.² Proponents, it is argued, should at least identify which strand of human dignity they are referring to (for example status, inner value, right, relating to treatment of human beings) so that we can make sense of its content and meaning. Considering human dignity is a pervasive idea in international human rights law and many constitutions,³ these criticisms seem unconvincing. In international humanitarian law, human dignity has been referred to as ‘the basic underpinning and ... the very *raison d’être* of international humanitarian law and human rights law ... in modern times it has become of such paramount importance as to permeate the whole body of international law.’⁴ It is given expression as an elementary consideration of humanity in the Martens Clause; a fundamental principle of customary international law protecting civilians and combatants in all circumstances not regulated by international law.⁵

² Michael Rosen, ‘Dignity: The Case Against’ in Christopher McCrudden (ed.), *Understanding Human Dignity* (OUP 2014) 143.

³ 1948 Universal Declaration of Human Rights, Preamble and Arts 1, 22, 23(3); 1966 International Covenant on Civil and Political Rights, Art. 10; 1966 International Covenant on Economic, Social and Cultural Rights, Art. 13; 1965 Convention on the Elimination of All Forms of Racial Discrimination, 1979 Convention on the Elimination of All Forms of Discrimination Against Women, 1984 Convention Against Torture and Other Cruel, Inhuman or Degrading Treatment or Punishment, Preamble; 1989 Convention on the Rights of the Child, Preamble and Arts 23, 28, 37, 39, 40; 2006 Convention on the Rights of Persons with Disabilities, Preamble and Arts 1, 3(a), 8(1)(a), 16(4), 24(1)(a), 25(d); 1978 Spanish Constitution, s 10(1); 1949 German Basic Law, arts 1(1) (as a duty), 79(3) (amendment to the duty is inadmissible); 1996 South African Constitution, s 1 (as a constitutional value), s 10 (as a right); see also Paolo Carozza, ‘Human dignity in constitutional adjudication’ in Tom Ginsberg and Rosalind Dixon (eds), *Comparative Constitutional Law* (Edward Elgar 2011).

⁴ *Prosecutor v. Furundžija*, Judgment, Case No IT-95-17/1-T, Trial Chamber (10 December 1998), para. 183.

⁵ Additional Protocol I to the Geneva Conventions of 12 August 1949 (1977) (‘API’), Art. 1(2); Hague Convention respecting the Laws and Customs of War on Land (‘Hague Convention IV’) (1907); Hague Convention with respect to the Laws and Customs of War on Land (‘Hague Convention II’) (1899).

Indeed, the ICJ has pointed to the significance of the Martens Clause as ‘an effective means of addressing rapid evolution of military technology’.⁶

In the 1949 Corfu Channel Case the ICJ referred to Albania’s international obligation to warn approaching British warships of the imminent danger posed by a minefield in Albanian territorial waters as based on ‘elementary considerations of humanity, even more exacting in peace than in war.’⁷ The 1996 ICJ Advisory Opinion on the *Legality of the Threat or Use of Nuclear Weapons* recognised that many rules of humanitarian law (for example, the distinction between civilians and combatants, prevention of unnecessary suffering, proportionality) are ‘so fundamental to the respect of the human person and “elementary considerations of humanity”’ that they have received wide acceptance through treaty ratification and customary international law.⁸ In the most recent case of *Costa Rica v. Nicaragua*, Judge Cançado Trindade reasserted the ‘autonomous legal regime’ of provisional measures, and how the principle of humanity has expanded their scope of protection beyond the inter-state dimension.⁹

These examples point to two interrelated aspects of human dignity: first, it is a status of human beings who have equal and inherent moral value and, second, rules (whether conceptualised as values, rights, or duties) flow from such a status relating to how human beings should be treated during war and in peace.¹⁰ Kant’s moral theory on ethical conduct provides a rationale and justification for rules based on human dignity as a status and as respectful treatment of human beings.

2. The Kantian Notion of Human Dignity

In his search for a moral theory on human conduct Kant put forward the idea of human dignity as a fundamental principle. In *Groundwork of the Metaphysics of Morals* (1785) he establishes the rationale and key elements of human dignity, and in

⁶ *Legality of the Threat or Use of Nuclear Weapons*, Advisory Opinion of 8 July 1996, ICJ Reports 1996, para. 78.

⁷ *Corfu Channel*, Merits Judgment, Judgment of April 9 1949, ICJ Reports 1949, 4, 22.

⁸ *Nuclear Weapons*, *supra* note 6, para. 79. See also paras. 92 and 95 on the principle of humanity.

⁹ *Certain Activities carried out by Nicaragua in the Border Area (Costa Rica v. Nicaragua)*, Judgment of 16 December 2015, ICJ Reports 2015, Separate Opinion of Judge Cançado Trindade.

¹⁰ Carozza expresses these as the ‘status claim’ and ‘normative principle’ under international human rights law: Paolo Carozza, ‘Human Dignity’ in Dinah Shelton (ed.), *The Oxford Handbook of International Human Rights Law* (OUP 2013), ch. 14.

his later work *The Metaphysics of Morals* (1797) tries to give practical illustration of its effect. Both works reveal Kant's understanding of human dignity as a special status conferred on humans from which certain rights and duties flow. This special status is based on the human capacity for rational thinking to create and abide by rules, and the capacity to identify what Kant refers to as 'ends'.¹¹ 'Ends' here means reasons or justifications for having the rules and abiding by them. Linked to rational thinking is 'autonomy of will', which refers to the human capacity to freely and rationally adopt and abide by rules.¹² These characteristics endow human beings with an overriding value known as human dignity. Dignity is an 'unconditional and incomparable worth' with 'intrinsic value' meaning it is not dependent on other factors for its existence, recognition, or respect. It is self-evident, priceless so cannot be replaced with an equivalent, and is 'the sole condition under which anything can be an end in itself'.¹³ Waldron criticises Kant for respecting something within a person rather than a person him or herself.¹⁴ But Kant's approach develops human dignity as a higher norm capable of applying to all humans, irrespective of recognition or not of their personhood, and entitling them to participate in law-making and governance with certain moral expectations.

Kant's notion of human dignity is inclusive and does not admit distinctions or exclusions on the basis of wrongdoing. If a person commits a wrong they do not lose their human dignity, which is something intrinsic and inherent to them. However, Kant's conception of what it takes to live in a moral society, referred to as 'the kingdom of ends', incorporates the need for state coercive force and punishment. Specifically in relation to wrongdoers who commit murder, he allows for the possibility of a death sentence, which seems at odds with human dignity. But because there is no equivalent to human dignity, once a life is taken through murder it can never be replaced so that the just thing to do is to take the life of the perpetrator. This is retributive punishment under strict conditions: there must first be a finding of

¹¹ Immanuel Kant, *The Moral Law: Kant's Groundwork of the Metaphysics of Morals* (H.J. Paton tr, Hutchinson & Co 1969) 90-91, paras. 64-66 [428-429].

¹² *Ibid.*, 101-102, paras. 87-88 [440].

¹³ Kant, *supra* note 11, at 90-91, paras. 65-67 [428-429]; 96-97, paras. 77-79 [435-436].

¹⁴ Jeremy Waldron, 'Dignity, Rank, and Rights' *The Tanner Lectures on Human Values* (University of California, Berkeley, 21-23 April 2009).

wrongdoing; the sentence must be judicially prescribed; and the wrongdoer must not be mistreated.¹⁵ Human dignity is upheld by punishing but not mistreating the wrongdoer for taking a life. Thus, human dignity is conceived as both a status and a higher norm governing how we treat each other. Below I will explain how and why autonomous weapons are contrary to the status and higher norm aspects of human dignity.

3. Autonomous Weapons and Human Dignity as a Status

Kant's reference to humanity as an objective end and humans as rational agents with autonomy of will helps explain how human dignity represents a status. Below I will explain the meaning and content of each of these elements and how autonomous weapons impact on them.

3.1 From humanity as an objective end to relative ends

What Kant refers to as humanity as an objective end is part of his process of establishing human dignity as a fundamental principle. Kant distinguishes 'relative ends' from 'objective ends'. Relative ends are values based on personal desires, wants, hopes, and ambitions. They are easily replaced and replaceable. Objective ends, however, cannot be replaced with an equivalent. They are reasons for morals governing human conduct which are capable of universalisation and valid for all rational beings. Objective ends are superior because they possess a particular moral value: dignity. Humanity as an objective end is expressed in Kant's maxim, 'Act in such a way that you always treat humanity, whether in your own person or in the person of any other, never simply as a means but always at the same time as an end.'¹⁶ What does it mean to treat someone as 'an end' rather than 'as a means'? Rational beings have intrinsic worth and a self-determining capacity to decide whether or not to do something. They are not mere objects or things to be manipulated, used or discarded on the basis of relative ends (for example personal wants, desires, hopes, and ambitions). Human dignity gives a person a reason for doing or not doing

¹⁵ Immanuel Kant, *The Metaphysics of Morals* (Mary Gregor tr. and ed., CUP 1996) 105-109.

¹⁶ Kant, *supra* note 11, at para. 67 [429].

something. That reason takes precedence over all others. It means setting moral and rational limits to the way we treat people in pursuit of relative ends.¹⁷

3.1.1 Treating human targets as mere objects

How does this relate to autonomous weapons? First, autonomous weapons are used for a relative end (that is, the desire to eliminate a human target in the hope of preventing harm to others). Relative ends, as we know from Kant's formulation, are lesser values capable of being replaced by an equivalent. This is not to say that preventing harm to others per se is a relative value. In fact, it is an objective end because it is something that all rational beings could freely and rationally agree to and abide by. But killing a human being in the hope that it will prevent further harm is insufficiently morally grounded to override human dignity and may be reckless if alternatives and consequences are not considered. Utilitarians may counter that balancing interests involves consideration of the greater good, which in this instance is to prevent harm to others.¹⁸ As Mill argued in relation to consequences, 'All action is for the sake of some end, and rules of action, it seems natural to suppose, must take their whole character and colour from the end to which they are subservient.'¹⁹ The double effect doctrine of harm, which differentiates between intended harm and unintended yet foreseeable harm, could also be used to justify action that results in unintended harm.²⁰ Indeed, consequentialist thinking and the utilitarian calculus are reflected in the proportionality principle under Article 51 API, requiring assessment of whether an attack is expected to cause excessive incidental loss of civilian life in relation to the concrete and direct military advantage anticipated.

However, such aspects of utilitarianism cannot overcome the problem of applying a quantitative assessment of life for prospective greater good that treats the humans sacrificed as mere objects, and creates a hierarchy of human dignity. In Germany, where the state has a constitutional duty to respect and protect human

¹⁷ For elaboration of Kant's humanity principle as an objective end representing human dignity, see Thomas E. Hill, Jr, *Dignity and Practical Reason in Kant's Moral Theory* (Cornell University Press 1992) 43–44.

¹⁸ Jeremy Bentham, *An Introduction to the Principles of Morals and Legislation* (Batoche Books, Kitchener [1781] 2000).

¹⁹ John Stuart Mill, *Utilitarianism* (Reprinted from *Fraser's Magazine*, 7th edn, Green and Co, published by "The Project Gutenberg eBook of *Utilitarianism*, by John Stuart Mill" [1879] 2004) 8-9.

²⁰ Fiona Woollard, *Doing and Allowing Harm* (OUP 2015).

dignity for all, such an approach was rejected by the Constitutional Court in 2006 when it declared void and unconstitutional aviation security legislation allowing the shooting down of hijacked planes. To sacrifice the lives of passengers and aircrew was to treat them as mere objects and call into question their quality and status as human beings with dignity.²¹ Thus, unless autonomous weapons can only be used to track and identify rather than eliminate a human target, they would extinguish a priceless and irreplaceable objective end possessed by all rational beings: human dignity. Perhaps the problem lies in utilitarianism setting the ‘end’ as preservation of human life rather than preservation of human dignity. Mill’s formulation that actions take ‘their whole character and colour from the end’ is another way of saying the ends justify the means. But if the ‘end’ is set as maximum preservation of human dignity in a broad sense (and not restricted to your own combatants) this cannot be achieved by autonomous weapons and would not satisfy rule-consequentialism or utilitarianism.

3.1.2 Undermining the existential reason for rules

Second, using autonomous weapons to extinguish life removes the reason for having morals in the first place: human dignity of rational beings with autonomy of will to apply rules of warfare. In doing so, a relative end is given priority over an objective end. Heyns warns:

it presents a very bleak picture of the international order if ethical norms are explicitly excluded from consideration. An approach that ignores ethical norms presents the spectre of an order that will find itself increasingly unsupported by the fundamental values of the people whose interests it is supposed to serve. Human rights norms such as the right to life and dignity have to be given contents in terms of ethical standards.²²

From a positivist or natural law theory approach, there is a basic existential reason for rules: to ensure states, peoples, and individuals can survive within the international

²¹ Bundesverfassungsgericht, February 15, 2006, 115, BVerfGE 118, paras. 121-124, available at www.bverfg.de/e/rs20060215_1bvr035705en.html (last visited 23 April 2017).

²² Christof Heyns, ‘Autonomous weapons systems and human rights law’ (Presentation made at the informal expert meeting organised by the state parties to the Convention on Certain Conventional Weapons 13-16 May 2014, Geneva, Switzerland) 8; Christof Heyns, ‘Autonomous weapons systems: living a dignified life and dying a dignified death’ in Nehal Bhuta, Susanne Beck, Robin Geiß, Hin-Yan Liu, Claus Kreß (eds), *Autonomous Weapons Systems Law, Ethics, Policy* (CUP 2016).

legal order.²³ But a rule that allows for life to be extinguished anywhere in the world by an autonomous weapon undermines the existential reason. Judge Weeramantry expanded on this point in relation to nuclear weapons in the *Legality of the Threat or Use of Nuclear Weapons*:

members of the international community have for the past three centuries been engaged in the task of formulating a set of rules and principles for the conduct of that society - the rules and principles we call international law. In so doing, they must ask themselves whether there is a place in that set of rules for a rule under which it would be legal, for whatever reason, to eliminate members of that community or, indeed, the entire community itself. Can the international community, which is governed by that rule, be considered to have given its acceptance to that rule, whatever be the approach of that community - positivist, natural law, or any other? Is the community of nations, to use Hart's expression, a "suicide club"?²⁴

3.1.3 Creating a hierarchy of human dignity

Third, without face-to-face killing certain humans are deemed more valuable and priceless than others, which creates a hierarchy of human dignity. This may appear to idealise or romanticise warfare as involving chivalrous, primitive physical contact between men.²⁵ And surely modern warfare increasingly involves distancing of combatants through the use of bombs and precision-guided missiles so that autonomous weapons are part of the same development? To clarify, face-to-face killing is not used here in the sense of caricaturing warfare as literally pitting men against each other through physical contact. Rather, it is intended to evoke the human essence of warfare; that there is human moral and legal agency in targeting and lethal force decisions and actions. Existing semi-autonomous weapons with some degree of autonomy in certain critical functions tend to be restricted in terms of types of operations, targets, and contexts (for example air defence weapon systems programmed to detect and shoot projectiles within close range of target base, border control defensive weapon systems programmed to detect and destroy incoming

²³ H. L. A. Hart, *The Concept of Law* (OUP 1961).

²⁴ *Nuclear Weapons*, *supra* note 6, Dissenting Opinion of Judge Weeramantry, 521.

²⁵ I am grateful to Marco Sassóli for raising this point during the ESIL conference. See also Marco Sassóli, 'Autonomous Weapons and International Humanitarian Law: Advantages, Open Technical Questions and Legal Issues to be Clarified' (2014) 90 *International Legal Studies* 308.

rockets, naval ship-based automatic machine guns programmed to detect and destroy incoming missiles).²⁶ Even with some removal of human involvement in relation to certain critical functions, these weapons remain dependent on human input and are designed to fire automatically when pre-determined parameters are detected. They are not designed to act independently in dynamic or changing situations.

So unlike other weapons, autonomous weapons, especially those with full autonomy in all the critical functions, present a more challenging impact by removing the human essence of warfare. Military personnel, remote pilots, commanders, programmers, and engineers are immune from rational and ethical decision-making to kill another human being and do not witness the consequences. By replacing the human combatant with a machine the combatant's human dignity is not only preserved but elevated above the human target. This can also be seen as a relative end in that it selfishly protects your own combatants from harm at all costs, including violating the fundamental principle of humanity as an objective end.²⁷

Characterising the protection of your own combatants as a relative and selfish end may offend utilitarian perspectives on the necessities and realities of warfare. Strawser distinguishes fully autonomous from semi-autonomous weapons, finding the former 'morally impermissible' but the latter, especially UAV, permissible due to a moral duty to protect soldiers who should not be put at unnecessary risk. UAV, he argues, do not violate the demands of justice, do not make the world worse, or expose your own combatants to potentially lethal risk unless incurring such a risk aids in the accomplishment of good in some way that cannot be gained via less risky means.²⁸ From a national interest and utilitarian perspective, this moral duty to protect soldiers sounds logical and sensible, but it fails to recognise that the inherent asymmetry in human dignity status introduced by autonomous weapons (whether fully or semi-autonomous) leads to insecurity and unpredictability in warfare which makes neither

²⁶ See, e.g., German NBS Mantis, available at <http://www.army-technology.com/projects/mantis/> (last visited 23 April 2017); Israeli Iron Dome, available at <http://www.army-technology.com/projects/irondomeairdefencemi/> (last visited 23 April 2017); American Phalanx, available at http://www.navy.mil/navydata/fact_print.asp?cid=2100&tid=487&ct=2&page=1 (last visited 23 April 2017).

²⁷ Hill pursues an interesting line of enquiry as to whether Kantian human dignity allows for this sort of hierarchy in relation to terrorists and hostage situations, Hill *supra* note 17, at ch. 10.

²⁸ Bradley Jay Strawser, 'Moral Predators: The Duty to Employ Uninhabited Aerial Vehicles' (2010) 9(4) *Journal of Military Ethics* 342-368.

the combatant nor the target safe. Even using a semi-autonomous weapon means a combatant is not in direct harm's way or at risk of losing their life. With a fully autonomous weapon a machine completely replaces the human combatant in all the critical functions. In either case, protection of the combatant relates to life, not dignity as status and respectful treatment. This in itself does not seem a valid reason for using autonomous weapons because potential loss of combatant lives in war is expected and an unavoidable risk (unless the state is negligent in preparing and equipping troops). In fact, replacing combatants with autonomous weapons undermines the former's dignity by not recognising their professional training and the military ethics of courage and respect for human targets, as reflected under international humanitarian law.²⁹ There would also appear to be no reason for having armies.

3.2 The cycle of irrationality and irrational agents

We have already established that Kant considers humanity as formed by rational beings with the capacity to create, amend, and abide by moral rules. Individuals engaged in immoral conduct are not excluded from humanity and, therefore, cannot lose their human dignity. Autonomy of will is key to Kant's conception of the rational being because it means individuals are not coerced to create, amend, and abide by moral rules.³⁰ Autonomy of will does not refer to the capacity to achieve personal objectives, which are relative ends. It is about freely and willingly accepting rules that achieve objective ends (for example preventing harm to humans in order to respect their human dignity). But the introduction of fully autonomous weapons with autonomy in all the critical functions actually makes us irrational agents who relinquish our autonomy of will. Humans are removed from the rational thinking process of when and how to use lethal force, and abdicate a key characteristic of humanity to a machine. This begs the question whether we need rationality at all if it can be so easily delegated to machines.

²⁹ Geoffrey Best, *Humanity in Warfare* (Columbia University Press 1980). See also Rain Liivoja, 'Chivalry Without A Horse: Military Honour and the Modern Law of Armed Conflict' (2012) 15 *Estonian National Defence College Proceedings* 75-100.

³⁰ Kant, *supra* note 11, at 107-109, paras. 97-101 [446-448].

‘Human central thinking activities’³¹ are critical during warfare and involve the ability to feel, think and evaluate, and the capacity to adhere to a value-based system in which violence is not the norm governing human relations. This uniquely identifies how humans engage in qualitative analysis through exercising judgment and reasoning. A combination of knowledge, experience, environment, and critical evaluation skills influence ‘human central thinking activities’ enabling difficult decisions on the extent and timing of force. Pre-programmed machines perform cost-effective and speedy peripheral processing activities based on quantitative analysis, repetitive actions, and sorting data. But they do not possess the human attributes to appraise a given situation, exercise judgment, refrain from taking action, or to limit harm. Stating that there will be human control over autonomous weapons is not enough to allay concerns about removing ‘human central thinking activities’ from the lethal force decision-making process. The type of human control is critical. A human-operated on/off switch to trigger an attack does not demonstrate the exercise of rational thinking. There is also the problem of automation bias in semi-autonomous weapons whereby the human operator accepts what the machine approves as legitimate targets. Sharkey refers to the need for ‘meaningful human control’, which means allowing human deliberation about a target before initiating an attack.³²

Without this rational capacity, do we then revert to a state of nature? Human targets are denied the status of rational agents with autonomy of will, and arbitrarily deemed irrational agents subject to extrajudicial killings or subhumans not worthy of human face-to-face contact. Remember that under the Kantian notion of human dignity immoral conduct does not lead to loss of human dignity, so no matter what the human target has or has not done they still have human dignity. By excluding the human target from human dignity on the basis of their alleged immoral conduct there is no opportunity to convince them of the validity of moral laws or to engage non-lethal methods. In fact, an opportunity is lost to build what Kant refers to as the ‘kingdom of ends’ in which rational beings create and abide by moral rules

³¹ Ozlem Ulgen, ‘Autonomous UAV and Removal of Human Central Thinking Activities: Implications for Legitimate Targeting, Proportionality, and Unnecessary Suffering’ (forthcoming) 1-45.

³² Noel Sharkey, ‘Towards a principle for the human supervisory control of robot weapons’ (2014) 2 (May-August) *Politica & Società* 1-16.

recognising human dignity. By violently ousting human targets for perceived irrational and immoral conduct, autonomous weapons perpetuate a cycle of irrationality in which humans become irrational agents.

4. Autonomous Weapons and Human Dignity as Respectful Treatment

Kant's approach to ethical conduct is rooted in rational beings with autonomy of will having an inclination towards respect for moral rules. This inclination derives from rationality and recognition of the intrinsic worth of human dignity. It is not based on self-interest or coercion. It follows from the status of human dignity that respecting the rights of others amounts to recognition of human dignity.³³ What are these rights? Respectful treatment of yourself and others is a manifestation of human dignity or humanity as an objective end. For example, human dignity resides in individuals taking care of their own moral worth through avoiding immoral conduct and constantly striving to move from a state of nature to an improved rightful or lawful condition. Individual morality is moderated by self-restraint and openness.³⁴ Too much self-restraint is contrary to human dignity (for example denial of basic human needs for some greater good). Too much openness in seeking personal pleasure at the expense of others is also contrary to human dignity (for example avarice, arrogance). As regards respecting others, Kant expresses this as a negative formulation. We restrain our words and deeds towards others and thereby respect their human dignity. Kant's writings on human value, state powers of punishment, and rights in war provide a basis for understanding human dignity as respectful treatment.

4.1 Mistreatment of rational beings and wrongdoers

Not mistreating human beings is Kant's negative formulation of the duty to respect human dignity in others. All humans, including wrongdoers, are rational beings with autonomy of will deserving respect for their human dignity. Recall that dignity means recognition of another's worth that has no price and cannot be exchanged for an

³³ Waldron refers to dignity as a 'status-concept' whereby humans with moral and/or legal standing control their own conduct, account for their actions, and demand that others recognise and accommodate such status. See Jeremy Waldron, 'How Law Protects Dignity' (2012) 71(1) Cambridge Law Journal, 201-202. See also Waldron, *supra* note 14.

³⁴ Kant, *supra* note 15, at 173-218.

equivalent. If we do not respect a wrongdoer's dignity or treat them less favourably we are judging them as worthless and, in Kant's terms, with contempt. For Kant a dangerous wrongdoer is no object of contempt and no less worthy of respect because he remains a human being even if his deeds are unworthy.³⁵ In relation to how to treat such a person, Kant refers to certain 'disgraceful punishments' that cannot be justified because they 'dishonour humanity itself ... [and] ... make a spectator blush with shame at belonging to the species that can be treated that way.'³⁶ Examples include quartering a man, having him torn by dogs, cutting off his nose and ears. These are severe acts against the physical integrity and dignity of the person which, read in conjunction with Kant's remarks about a judicially prescribed death sentence without mistreatment, provide illustrations of the duty not to mistreat humans. More subtle illustrations of mistreatment, referred to as 'vices', include arrogance, defamation, and ridicule.³⁷

4.1.1 'Outrages upon personal dignity' and inhumane treatment

Kant's notion of human dignity conceptualises the generic category of 'wrongdoers' to help us understand that even if a person is suspected of wrongdoing or has done wrong, or is an enemy combatant, they are still entitled to status and certain treatment. Kant's 'disgraceful punishments' are today transposed into international humanitarian law through prohibition of certain acts and forms of conduct. Common Article 3 of the Geneva Conventions provides fundamental guarantees (applicable to both non-international and international armed conflicts) that civilians and hors de combat 'shall in all circumstances be treated humanely'.³⁸ Article 3(1)(a) prohibits violence to life and person, in particular murder of all kinds, mutilation, cruel treatment and torture. Article 3(1)(c) prohibits 'outrages upon personal dignity, in particular humiliating and degrading treatment'. The Elements of Crimes for the International Criminal Court defines 'outrages upon personal dignity' as acts which humiliate,

³⁵ Kant, *supra* note 15, at 210.

³⁶ *Ibid.*

³⁷ Kant, *supra* note 15, at 211-213.

³⁸ The majority decision in *Military and Paramilitary Activities in and against Nicaragua (Nicaragua v. USA)*, Merits Judgment of 27 June 1986, ICJ Reports 1986, 14 held that Common Art. 3 expresses 'minimum rules applicable to international and non-international conflicts' (para. 219), and these rules reflect 'elementary considerations of humanity' (para. 218).

degrade, or otherwise violate the dignity of a person to such a degree ‘as to be generally recognised as an outrage upon personal dignity’.³⁹

Common Article 3 fundamental guarantees are also provided for enemy combatants under Articles 1(2) and 75 of API. Enemy combatants are afforded protection under ‘the principles of international law derived from established custom, from the principles of humanity and from the dictates of public conscience’; and if they do not benefit from more favourable treatment under the Geneva Conventions or API, they must be ‘treated humanely in all circumstances’. The law’s moral basis derives from ‘principles of humanity’ and ‘the dictates of public conscience’ which, although not defined, are intended to overcome any ambiguities or uncertainties by anchoring the law back to what would be in the interest of humanity. This moral basis prevents the assumption that something which is not prohibited in law is therefore permissible, and applies regardless of developments in weapons technology.⁴⁰ It has normative force to provide additional protection by appropriately controlling military behaviour.⁴¹

These provisions establish obligations to take account of others’ interests, including the human dignity of enemy combatants. Use of autonomous weapons to kill ‘wrongdoer’ human targets completely bypasses such obligations and represents a modern-day example of Kant’s ‘disgraceful punishments’ amounting to ‘outrages upon personal dignity’. The human target is treated as an inanimate object without any interests; easily removed and destroyed by a faceless and emotionless machine. No value is placed on the life taken. No ‘human central thinking activities’ are involved in the interpretation and application of international humanitarian law on prevention of unnecessary suffering, taking precautionary measures, and assessing

³⁹ Elements of Crimes for the International Criminal Court, Art. 8(2)(b)(xxi) (war crime of outrages upon personal dignity covering all persons including the dead); Art. 8(2)(c)(ii) (war crime of outrages upon personal dignity of hors de combat, civilians, medical personnel, or religious personnel).

⁴⁰ Yves Sandoz, Christophe Swinarski and Bruno Zimmermann (eds), ICRC Commentary on the Additional Protocols of 8 June 1977 to the Geneva Conventions of 12 August 1949 (ICRC 1987), paras. 55-56.

⁴¹ *Nuclear Weapons*, *supra* note 6, Dissenting Opinion of Judge Shahabudden, 405-409.

proportionality. The lack of human discretion in these decisions violates Articles 35, 51, 57 API.⁴²

There is currently no prohibition on the use or development of autonomous weapons but this does not make them permissible when judged against human dignity as a principle of humanity. Autonomous weapons would devalue humanity by treating humans as disposable inanimate objects rather than ends with intrinsic value and rational thinking capacity. All individuals targeted and killed by such weapons are entitled to respect for their human dignity. Whether or not they are designated enemy combatants or terrorists, they have rational capacity, possess a moral value of dignity which cannot be replaced by an equivalent, and cannot lose such status through immoral acts. If an autonomous weapon is capable of causing unnecessary suffering in the human target this would constitute mistreatment. For example, certain types of Hellfire missiles used on UAV cause burning in targets and incineration of bodies.⁴³ The AGM-114N MAC ('metal augmented charge') variant uses a thermobaric warhead that can 'suck the air out of a cave, collapse a building, or produce an astoundingly large blast radius out in the open.'⁴⁴ It contains a 'fluorinated aluminium powder layered between the warhead casing and the PBXN-112 explosive fill. When the PBXN-112 detonates, the aluminium mixture is dispersed and rapidly burns. The resultant sustained high pressure is extremely effective against enemy personnel and structures.'⁴⁵

4.1.2 Does it matter whether mistreatment comes from man or machine?

⁴² Eliav Lieblch and Eyal Benvenisti, 'The obligation to exercise discretion in warfare: why autonomous weapons systems are unlawful' in Bhuta (et al.) *supra* note 22, argue that autonomous weapons systems violate the duty to exercise discretion under international humanitarian law because they have pre-determined decision-making capability which does not respect the individual by considering their case/position carefully and exercising discretion where necessary.

⁴³ See e.g., American UAV strike of 23 January 2013 killing four individuals, including two civilians: Joseph Cox, 'The Yemeni Man Suing BT for America's Deadly Drone Attacks' Vice News (23 May 2014) available at www.vice.com/en_uk/read/the-yemeni-man-suing-bt-for-american-drone-strikes (last visited 23 April 2017); Glenn Greenwald, 'Burning Victims to Death: Still a Common Practice' The Intercept (4 February 2015) available at <https://theintercept.com/2015/02/04/burning-victims-death-still-common-practice/> (last visited 23 April 2017).

⁴⁴ 'US Hellfire Missile Orders, FY 2011-2017' available at www.defenseindustrydaily.com/us-hellfire-missile-orders-fy-2011-2014-07019/ (last visited 23 April 2017).

⁴⁵ 'AGM-114N Metal Augmented Charge (MAC) Thermobaric Hellfire' available at www.globalsecurity.org/military/systems/munitions/agm-114n.htm (last visited 23 April 2017).

It may be argued that international humanitarian law allows use of lethal force against an enemy so that death resulting from use of autonomous weapons is not unlawful per se. But this avoids moral and legal considerations of methods and means of warfare, which stand at the heart of human dignity as respectful treatment. To say that human targets are indifferent as to whether they are killed by autonomous weapons or soldiers undermines human dignity in the person and runs contrary to evidence of the effects and repercussions of American UAV strikes in Pakistan and Yemen.⁴⁶ Apart from causing civilian casualties, UAV strikes have caused loss of livelihood due to fear of venturing outside, and severe psychological harm officially diagnosed as PTSD.⁴⁷ ‘Decapitation strikes’ intended to weaken the organisational capability of al-Qaeda and the Taliban by removing key players or leaders have not achieved that objective, and UAV strikes in Pakistan have fuelled recruitment into militant organisations and solidified resistance against the Pakistani State.⁴⁸ Victims’ accounts of targeted killings in Yemen reveal extreme physical, psychological and economic harm: targeted vehicles continue burning with victims inside; clothes fused to survivors’ skin; skin burned off; local population living in fear and terror from hearing planes; women suffering miscarriages; children frightened to go outside; dependents of individuals killed unable to support themselves economically; local population suffering shock after strikes; inability to sustain a living from the land due to fear of being outside.⁴⁹

Cases of combatants committing war crimes amounting to ‘outrages upon personal dignity’ may appear to bolster the argument supporting use of autonomous weapons (that is, the latter will act more rationally and be less prone to human flaws

⁴⁶ Dieter Birnbacher, ‘Are autonomous weapon systems a threat to human dignity?’ in Bhuta (et al.) *supra* note 22, asserts that victims whose human dignity is at stake are indifferent as to whether they are killed by manned or unmanned weapons, although concedes that use of autonomous weapons systems in certain circumstances (e.g. causing unrelieved mental pain for civilian victims, severely restricting their freedom, and potentially treating victims as means to an end) potentially violates human dignity.

⁴⁷ Stanford International Human Rights and Conflict Resolution Clinic and Global Justice Clinic, *Living under drones: death, injury, and trauma to civilians from US drone practices in Pakistan* (NYU School of Law, September, 2012), ch. 3; UK Defence Committee, *Written evidence from the All Party Parliamentary Group on Drones (APPG) (HC 2013-14)* para. 24 (Dr Peter Schaapveld’s, forensic psychologist, evidence on drones in Yemen).

⁴⁸ Jenna Jordan, ‘Attacking the Leader, Missing the Mark: Why Terrorist Groups Survive Decapitation Strikes’ (Spring 2014) 38(4) *International Security* 7; Saira Yamin and Salma Malik, *Mapping Conflict Trends in Pakistan* (United States Institute of Peace, *Peaceworks* No. 93, February 2014) 7.

⁴⁹ *Reprieve Complaint to the UK National Contact Point under the Specific Instance Procedure of the OECD Guidelines for Multinational Enterprises in respect of BT plc* (15 July 2013) 4-5.

leading to atrocities). The ‘irrational soldier’ argument maintains that soldiers are susceptible to emotions and unpredictability which can be eliminated by use of autonomous weapons.⁵⁰ Human emotions and negative human characteristics (for example susceptibility to fatigue and capacity for revenge) are said to debilitate soldiers’ performance and ethical judgment to the extent of enabling commission of war crimes and atrocities. Replacing soldiers with autonomous weapons will somehow guarantee rational thinking which does not lead to violations of international humanitarian law or war crimes.

But such an argument makes questionable assumptions about human and robotic rational thinking capacities and conduct in warfare, disregarding human agency in the creation and failures of autonomous weapons. It assumes humans have no capacity for preventing unethical conduct and that machines will act more ethically than humans. There are many different reasons why combatants commit war crimes, not necessarily related to inherent human flaws.⁵¹ Human emotions, as part of ‘human central thinking activities’, play a vital role in navigating complex social environments in combat, especially where it is necessary to perceive and interpret human behaviour (for example children playing ball rather than throwing a hand grenade, someone running with a stick rather than a gun, a young man of military age in the vicinity of an attack).⁵² The judgment, reasoning, and discretion exercised by a human cannot be performed by a machine. Far from advocating replacement of human combatants with machines, war crimes cases serve as barometers of public conscience on acceptable conduct in warfare. They promote human dignity by recognising that only human action justifies lethal force and, therefore, requires human accountability and responsibility.

⁵⁰ Ronald Arkin, *Governing Lethal Behaviour in Autonomous Robots* (CRC Press 2009); Ronald Arkin, ‘Lethal Autonomous Systems and the Plight of the Non-combatant’ (work was supported in part by the U.S. Army Research Office under Contract #W911NF-06-1-0252).

⁵¹ Daniel Munoz-Rojas and Jean-Jacques Frésard, *The Roots of Behaviour: Understanding and Preventing IHL Violations* (Geneva: International Committee of the Red Cross, 2004), found that advanced technologies which permit killing at a distance or on the computer screen prevent the activation of neuro-psychological mechanisms which render the act of killing difficult, and violations are usually coupled with moral disengagement. See also Hugh Gusterson, *Drone: Remote Control Warfare* (MIT Press 2016) 66-77, on the impact of UAV ‘voyeuristic technology’ and ‘remote narrativization’.

⁵² Marcello Guarini and Paul Bello, ‘Robotic Warfare: Some Challenges in Moving from Noncivilian to Civilian Theaters’ in Patrick Lin, Keith Abney and George Bekey (eds), *Robot Ethics: the ethical and social implications of robotics* (MIT Press 2012).

In *R v. Blackman* a British Acting Colour Sergeant in the Royal Marines, in command of a group of the Royal Marines serving as part of the British Armed Forces in Afghanistan, was found guilty of murder and sentenced to life imprisonment for shooting in the chest a seriously wounded Afghan insurgent. The insurgent was entitled to be treated with dignity, respect, and humanity, yet the non-commissioned officer treated him ‘with contempt and murdered him in cold blood’.⁵³ The Sergeant had failed to ensure medical assistance was quickly provided, allowed soldiers under his command to manhandle the wounded insurgent causing him additional pain, ordered those providing first aid to stop, and waited for the military surveillance helicopter to be out of sight before shooting the insurgent. The Sergeant was filmed saying, ‘There you are, shuffle off this mortal coil, you cunt ... It’s nothing you wouldn’t do to us,’ then turned to his fellow soldiers and stated, ‘Obviously this doesn’t go anywhere, fellas ... I’ve just broke [*sic*] the Geneva Convention,’ while the insurgent continued to writhe as these remarks were made.⁵⁴ Whilst mitigating factors pointed to human flaws in combat (for example the effect of fellow soldiers’ injuries and deaths, combat stress) these were not extraordinary or unexpected risks to sufficiently displace military ethics and respect for humanity:

... thousands of other Service personnel have experienced the same or similar stresses. They exercised self discipline and acted properly and humanely; you did not. ... while this sort of offence is extremely rare, if not unique, those Service personnel who commit crimes of murder, or other war crimes or crimes against humanity while on operations will be dealt with severely. This is a message of deterrence but it is

⁵³ *R v. Sergeant Alexander Wayne Blackman* (UK Military Court, Bulford, Judge Advocate General Jeff Blackett, Sentencing Remarks, 6 December 2013) Case Reference: 2012CM00442, available at www.judiciary.gov.uk/wp-content/uploads/JCO/Documents/Judgments/r-v-blackman-marine-a-sentencing+remarks.pdf (last visited 23 April 2017).

⁵⁴ *R v. Alexander Wayne Blackman* [2017] EWCA Crim 190 (15 March 2017), para. 22. See also Steven Morris, ‘Judges Allow Partial Release of British Marine Shooting Video’, *The Guardian* (1 February 2017), available at www.theguardian.com/uk-news/2017/feb/01/judges-allow-partial-release-of-british-marine-shooting-video (last visited 23 April 2017). On appeal his conviction was substituted for manslaughter on grounds of diminished responsibility due to new evidence adduced that he had suffered from adjustment disorder, see *R v. Alexander Wayne Blackman* [2017] EWCA Crim 325 (28 March 2017). His sentence was reduced to seven years which, allowing for time already spent in prison, resulted in him serving three and a half years in prison and being released in April 2017. His previous dismissal with disgrace was substituted for dismissal. For criticism about the Court Martial Appeal Court’s application of the defence of diminished responsibility and sentencing leniency, particularly given the defendant’s manifest premeditation and awareness of his actions, see Kevin Jon Heller, ‘Bad Criminal Law in the Alexander Blackman Case (With Addendum)’ *Opinio Juris* (31 March 2017) available at http://opiniojuris.org/2017/03/31/bad-criminal-law-in-the-alexander-blackman-case/?utm_source=feedburner&utm_medium=email&utm_campaign=Feed%3A+opiniojurisfeed+%28Opinio+Juris%29 (last visited 23 April 2017).

also to reassure the international community that allegations of serious crime will be dealt with transparently and appropriately.⁵⁵

Although murder is separate from and not an ‘outrage upon personal dignity’,⁵⁶ the sergeant’s acts and omissions preceding the killing can be characterised as ‘animated by contempt for the human dignity of another person’⁵⁷ and, therefore, falling into the particular category of inhumane treatment. This case makes clear the importance of interaction and interrelatedness between warring parties in the application of international humanitarian law. The court emphasised that military personnel acting with brutality and savagery lose the support and confidence of those they seek to protect, and provoke the enemy to act more brutally in retribution or reprisal. With use of autonomous weapons there is no interaction and interrelatedness between warring parties, which creates a dangerous human accountability and responsibility gap.

4.1.3 Mistreatment of the deceased

‘Outrages upon personal dignity’ and inhumane acts may be committed against the dead in warfare, and the Geneva Conventions establish extensive state obligations to search for the dead and prevent their being despoiled or ill-treated.⁵⁸ International criminal tribunal cases provide examples of the types of acts and conduct that constitute outrages against the personal dignity of the deceased, and consideration must be given as to how autonomous weapons may affect this aspect of human dignity.

In the Trial of Max Schmid, a German medical officer was found guilty of wilfully, deliberately and wrongfully mutilating the deceased body of a US serviceman, and sentenced to ten years imprisonment.⁵⁹ In *Prosecutor v. Niyitegeka* the Minister of Information in the Rwandan Interim Government was convicted,

⁵⁵ *R v. Sergeant Alexander Wayne Blackman*, *supra* note 53.

⁵⁶ *Prosecutor v. Kvočka and others*, Judgment, Case No IT-98-30/1-T, Trial Chamber (2 November 2001), para. 172.

⁵⁷ *Prosecutor v. Aleksovski*, Judgment, Case No IT-95-14/1-T, Trial Chamber (25 June 1999), paras. 54-56.

⁵⁸ Art. 15 Geneva Convention I of 12 August 1949; Art. 18 Geneva Convention II of 12 August 1949; Arts. 13, 120-121 Geneva Convention III of 12 August 1949; Art. 16 Geneva Convention IV of 12 August 1949; Art. 34 API.

⁵⁹ *Trial of Max Schmid*, Case No 82, UN War Crimes Commission Law Reports of Trials of Criminals, 19 May 1947, 151-152.

among other crimes, of other inhumane acts as crimes against humanity. He was jubilant at the capture of a prominent Tutsi and rejoiced when he was killed, decapitated, castrated, his skull pierced through the ears with a spike, and his genitals hung on a spike for the public to see. Niyitegeka's jubilation, especially in light of his leadership role in the attack, supported and encouraged the attackers and thereby aided and abetted the commission of crimes. He subsequently ordered men to undress the deceased body of a Tutsi woman and insert a sharp piece of wood into her genitalia. The Trial Chamber considered the order an aggravating factor for its 'cruel and insensitive disregard for human life and dignity' and found that both incidents 'would cause mental suffering to civilians, in particular, Tutsi civilians, and constitute a serious attack on the human dignity of the Tutsi community as a whole'.⁶⁰

These criminal acts cannot simply be explained as consequences of combat because they go beyond what is necessary to achieve a military objective and offend humanity. Although Kantian human dignity is predicated on the individual, *Prosecutor v. Niyitegeka* suggests there is a communitarian aspect to human dignity concerning mistreatment of deceased human targets. 'Civilians', 'Tutsi civilians', 'the Tutsi community as a whole' were all deemed to be affected as a group of protected persons, as an ethnic group, and as part of humankind. 'Collective human dignity' is important in communities where humaneness and personhood are achieved through association with others. Social honour, group moral standing, and the capacity to form communal relations are aspects of the 'ubuntu' tradition in sub-Saharan Africa.⁶¹ To the extent that 'ubuntu' regards dignity as conceived and maintained by communities it offers a perspective on what it means to be part of humankind. It would, for example, be wrong to punish or otherwise harm an innocent person, even when it would save more innocent lives, because this fails to treat the person in accordance with the way he has exercised his capacity for community. But this communitarian aspect has a problematic potential to exclude individuals on the basis that they are not 'innocents' or do not meet communitarian standards. What happens to the enemy

⁶⁰ *Prosecutor v. Niyitegeka*, Judgment, Case No ICTR-96-14-T, Trial Chamber (16 May 2003), paras. 465 and 499.

⁶¹ On 'collective human dignity' see Mica Werner, 'Individual and collective dignity' in Marcus Düwell, Jens Braarvig, and Roger Brownsword (eds), *The Cambridge Handbook of Human Dignity* (CUP 2014), ch. 35; Thaddeus Metz, 'Dignity in the ubuntu tradition' in Marcus Düwell (et al.) *ibid.*, ch. 32.

combatant who is not part of any community or is unable to form communal relations? Does he lack personhood and human dignity? Would this then justify using autonomous weapons against him? Kantian human dignity avoids this dilemma by advocating humane treatment of wrongdoers and enemy combatants.

Beyond this particular problematic potential, the communitarian aspect offers a way of understanding how autonomous weapons offend the ‘collective human dignity’ of mankind. Lack of face-to-face contact (understood as human moral and legal agency) between human combatants and human targets renders state obligations to search for and prevent mistreatment of the dead redundant. The weapon is programmed to administer lethal force without concerns about how the target dies or what happens after death. There is no obligation for the weapon to search for the dead or prevent mistreatment, which itself could be considered a form of mistreatment. Death caused by conventional combat (that is, human combatants in situ) whether on land, at sea, or in the air appears to be accorded greater protection against outrages upon personal dignity than death by autonomous weapons.

4.2 Preconditions for punishment of wrongdoers and treatment of enemy combatants

Some ambiguities exist in Kant’s formulation of dignity which may raise difficulties in its application to those targeted and killed by autonomous weapons. What if those targeted have killed humans? Do they lose rational capacity? Should they be afforded dignity? There are two possible answers here.

The first focuses on rational capacity as a potential rather than actual human characteristic and, therefore, dignity cannot be lost by committing immoral acts. This corresponds to the principle of equality based on innate humanity of all persons so that if you kill the person committing an immoral act, you kill yourself.⁶² The second answer provides a potential exception under punishment of such individuals. Kant regards retributive punishment, specifically the death penalty for murderers, as a matter of justice. Life and death are not the same and there is no substitute for taking a life other than death. But this exception has preconditions: punishment and sentence

⁶² Kant, *supra* note 15, at 105.

must be imposed by a judge, and even if the wrongdoer is facing an imminent death sentence ‘he must still be freed from any mistreatment that could make the humanity in the person suffering it into something abominable.’⁶³ Kant gives an example of mistreatment as dangerous physical experiments on a murderer for the greater medical good. These can never be consented to by the murderer or society because they are contrary to the murderer’s human dignity and ‘justice ceases to be justice if it can be bought for any price whatsoever.’⁶⁴

Kant’s non-mistreatment precondition is pervasive in today’s international humanitarian law. Even in situations where security or repressive measures are necessary against certain individuals, the dictates of humanity require that the law provides protection from mistreatment in order to preserve human dignity.⁶⁵ For example, Articles 13 and 14 Geneva Convention III and Article 11(1) API require prisoners of war to be treated humanely at all times; not be subjected to physical mutilation or to medical or scientific experiments, even with their consent; to be protected from acts of violence or intimidation and against insults and public curiosity; and to be entitled in all circumstances to ‘respect for their persons and their honour’. Thus, even if those targeted by autonomous weapons have killed human beings, they are still entitled to humane treatment.

The preconditions (that is, judicial punishment and sentence, no mistreatment) keep human dignity intact by not legitimising cruel and arbitrary treatment of wrongdoers as human outcasts without any moral rights. Human dignity is innate, priceless, and an objective end in itself. But the difference with autonomous weapons is that there is no due process to determine guilt or innocence. There is no prior determination of punishability and a judge is not imposing a sentence. Although lawful killing of enemy combatants in armed conflict generally does not require a prior judgment because it is not about punishment, autonomous weapons as a means of killing represent a form of punishment without preconditions because they undermine human dignity and deny the possibility of interaction and interrelatedness

⁶³ Kant, *supra* note 15, at 106.

⁶⁴ Kant, *supra* note 15, at 105.

⁶⁵ Jean S. Pictet (ed.), ICRC Commentary on Geneva Convention III 12 August 1949 (ICRC 1960) 140.

between warring parties. By targeting and attacking the wrongdoer the autonomous weapon is a means of avoiding judicial pronouncement and authorisation of punishment. The human target is treated as a means to an end without human dignity and subhuman treatment of efficient disposal is justified.

The non-existence of preconditions is contrary to Kant's ideal state of the 'kingdom of ends' (a commonwealth of persons who legislate universal laws that are rational and based on humanity as an end in itself). Kant is not advocating a world superstate but recognises the need for some form of state apparatus to enable legislating in this kingdom of ends. State apparatus necessarily includes coercive and punishment powers. Those who go against the moral rules can be punished, but not in a way that mistreats them or is contrary to human dignity. There must be an opportunity for the wrongdoer to avoid the punishment, which in any case must be judicially prescribed and administered. As we have already seen, Kant considers the death penalty a legitimate sentence for murderers but only under judicially prescribed conditions and only if the murderer is not mistreated in any other way.⁶⁶

4.3 Limitations on methods and means of warfare

Kant's idea of individual and state morality is based on a trajectory from a state of nature to a rightful or lawful condition. It forms the basis for his views on rights in war. Kant describes war as 'barbaric' and to be expected while states remain in a state of nature.⁶⁷ A state of rightfulness would involve states voluntarily coming together in a congress to uphold perpetual peace. Conceding that the state of nature will involve war, he then discusses rights in war. Where there is an 'unjust enemy' states are entitled to unite against and deprive the enemy state of its power. An 'unjust enemy' is one 'whose publicly expressed will (whether by word or deed) reveals a maxim by which, if it were made a universal rule, any condition of peace among nations would be impossible and, instead, a state of nature would be perpetuated.'⁶⁸

⁶⁶ Kant, *supra* note 15, at 105-109.

⁶⁷ Kant, *supra* note 15, at 118-120, paras. [6:349-6.351].

⁶⁸ Kant, *supra* note 15, at 119, para. [6:349].

4.3.1 Kantian ‘just war’ impact on jus in bello

Kant’s reference to just war theory is somewhat problematic in merging jus ad bellum and jus in bello considerations. But it is understandable for the period in which he was writing. Early modern writers of the 16th to 18th centuries were busy expanding Aquinas’ classical just war theory, which was conceived as a limited enforcement measure requiring three conditions: a sovereign authority to wage war; a just cause (broadly defined as a prior or threatening injury by the enemy); and an intention to do justice and attain a just peace.⁶⁹ By the 18th century it encompassed all-out war among sovereigns and legitimised broadly defensive action which did not necessarily limit violence against an enemy.⁷⁰ Vitoria recognised that both warring parties could legitimately claim to be fighting a just cause, and therefore had the right to derive benefit from jus in bello. Gentili considered war lawful if it was waged by a sovereign and formally declared. Grotius introduced the concept of ‘state of war’ which meant jus in bello applied to belligerents and jus in pace (that is, normal laws) applied in peacetime. Unlike these writers, Kant, even with his opposition to war, saw the necessary evil in having a limited form of self-help as conceived under classical just war theory. But he seems to be sabotaging it from within by setting limitations on methods and means of warfare to include proportionality in the conduct of hostilities, and humane treatment of enemy combatants. It is at this juncture that Kant’s understanding of jus in bello proves significant for autonomous weapons.

Modern just war theory has nothing to say on methods and means of warfare, but requires us to conceive of the slippery-slope ‘just cause’ war.⁷¹ It is highly contentious and subjective as to what constitutes a ‘just cause’ and all parties to a conflict may conceivably claim they are pursuing a just cause that justifies use of autonomous weapons. The utilitarian argument for protecting your own soldiers fails to recognise the need for interaction and interrelatedness in warfare in order for some

⁶⁹ Randall Lesaffer, ‘Too much history. From war as sanction to the sanctioning of war’ in Mark Weller (ed.), *The Oxford Handbook of the Use of Force in International Law* (OUP 2015); Mary Ellen O’Connell, ‘Peace and War’ in Bardo Fassbender and Anne Peters (eds), *The Oxford Handbook of the History of International Law* (OUP 2014), ch. 11.

⁷⁰ Lesaffer, *ibid.*

⁷¹ Michael Walzer, *Just and Unjust Wars: A Moral Argument With Historical Illustrations* (Basic Books 1977). For example, Strawser, *supra* note 28, argues for a duty to use UAV in order to prevent harm to soldiers fighting a just war: ‘for any just action taken by a given military, if it is possible for the military to use UAV platforms in place of inhabited aerial vehicles without a significant loss of capability, then that military has an ethical obligation to do so’, at 346.

basic common rules to apply to all parties and limit harm. A dual-harm prevention approach is necessary in the conduct of warfare; one that takes account of combatants and targets so as to lay foundations for common understanding and de-escalation of war. Kantian cosmopolitan ethics places human beings at the centre of norms and laws governing warfare so that irrespective of combatant or target status, both are entitled to human dignity. Seeking to justify any methods and means of warfare on the basis of a just cause disregards the impact on humans.

4.3.2 Restraint to preserve human dignity and conditions for peace

Kant wrote in a period of rapid state formation with states as the primary global force for power, but his definition of ‘unjust enemy’ could extend to non-state actors with transnational effects. Thus, whatever the political or religious motives of Islamist suicide bombings, such arbitrary attacks are not capable of universalisation as a moral rule because they guarantee mutual self-destruction, perpetuate war, and are fundamentally opposed to humanity in the Kantian sense of rational beings with autonomy of will. This being so, such attacks threaten state freedoms so that states are entitled to unite and deprive Islamist groups of power. But even when discussing permissible actions against unjust enemies Kant maintains a perspective of rational and proportionate conduct aimed at preserving peace and respecting human dignity. Thus, states are not entitled to divide the enemy’s territory among themselves or to eliminate the state because this would be an ‘injustice against its people.’ Distinguishing an unjust enemy state from its people enables prohibiting acquisition of land by force and respect for human dignity by treating humanity as an objective end and not simply as a means.

A state that is injured by the actions of an unjust enemy cannot resort to any means of self-defence but ‘those means that are allowable to any degree that it is able to, in order to maintain what belongs to it.’⁷² Kant offers no further explanation as to why there should be restraint on methods and means of warfare, but the same reasoning used to justify forms of state punishment against wrongdoers can apply here. Resorting to any means of retaliation, irrespective of its impact on peoples or the

⁷² Kant, *supra* note 15, at 119, para. [6:349].

degree to which it is necessary, would continue the state of nature and deny the existence of rational beings with autonomy of will. He identifies poisoning, using your own citizens as spies, and using your own citizens or foreigners as assassins or poisoners as impermissible methods and means of self-defence because these are ‘underhanded means as would destroy the trust requisite to establishing a lasting peace in the future.’⁷³ So Kant’s rationale for restraint on methods and means of warfare is based on preserving human dignity as well as ensuring conditions exist for perpetual peace among states. At its core is interaction and interrelatedness between rational beings and wrongdoers, and injured states and unjust enemies.

Such interaction and interrelatedness is evident in today’s international humanitarian law. For example, poison and poisoned weapons are prohibited under Article 23(a) of the Regulations annexed to the 1899 Hague Convention II and the 1907 Hague Convention IV and their use constitutes a war crime under Article 8(2)(b) (xvii) of the 1998 Rome Statute of the International Criminal Court. A spy captured out of the uniform of his armed forces whilst engaging in espionage loses combatant status and POW rights, but must still be treated humanely according to the fundamental guarantees under Article 75 API.⁷⁴

Restraint in action towards human targets is reflected in Articles 35 and 36 API, the principle of proportionality, and prevention of unnecessary suffering. If we accept that humanitarian law exists to regulate conduct in warfare in order to limit unnecessary suffering then it certainly is not controversial to expect less lethal methods and means, where possible, to be used. Capturing enemy combatants where feasible is an example and arguably a duty derived from custom and treaty law.⁷⁵ Current use of remotely operated UAVs involves pursuing individuals to their death without prior criminal investigation, due process, or attempts at their capture. The same lethal capability and lack of non-lethal options would exist with autonomous weapons, especially those with autonomy in all the critical functions. The elimination

⁷³ Kant, *supra* note 15, at 117, para. [6:347].

⁷⁴ Regulations annexed to the 1899 Hague Convention II and the 1907 Hague Convention IV, Art. 29; Art. 46 API.

⁷⁵ See, e.g., ICRC Interpretive Guidance on the Notion of Direct Participation in Hostilities under International Humanitarian Law (Geneva, Switzerland, May 2009); Ryan Goodman, ‘The Power to Kill or Capture Enemy Combatants’ (2013) 24(3) EJIL 819-853. For an opposite view see, W. Hays Parks, ‘Part IX of the ICRC ‘Direct Participation in Hostilities’ Study: No Mandate, No Expertise, and Legally Incorrect’ (2010) 42 NYU J Int’l L & Politics 794; Michael Schmitt., ‘Wound, Capture, or Kill: A Reply to Ryan Goodman’s ‘The Power to Kill or Capture Enemy Combatants’ (2013) 24(3) EJIL 855-861.

capability of autonomous weapons creates injustice against targeted individuals, physical and psychological collateral damage to communities in targeted areas, and perpetuates a state of nature.

But could restraint needs be overridden by utilitarian considerations of the costly resourcing and financing of conventional warfare and, therefore, justify replacing combatants with efficient machines? It is claimed that time and money can be saved by investing in autonomous weapons technology to engage in combat situations with precision and efficiency.⁷⁶ Autonomy in unmanned air systems, for example, is seen as a means to decrease the number of personnel needed to operate them.⁷⁷ It may require only one person to control multiple unmanned systems with automated processing and analysis of information. Autonomous weapons might also be used to substitute or expand existing ground forces.⁷⁸ These practical, cost-benefit reasons for supporting autonomous weapons sit rather uncomfortably with more pressing concerns about legitimate targeting, proportionality, and preventing unnecessary suffering.

Cost-benefit reasoning was used to justify the fire-bombing of German and Japanese cities during the Second World War and, eventually, the nuclear bombing of Hiroshima. In a critical essay about nuclear warfare Rawls rejects this reasoning because it ‘justifies too much, too easily, and provides a way for a dominant power to quiet any moral worries that may arise. If the principles of war are put forward at that time, they easily become so many more considerations to be balanced in the scales.’⁷⁹ In the same way, cost-benefit reasoning in the use of autonomous weapons ‘too easily’ provides justification and a way to ‘quiet any moral worries’. Indeed, the cost-benefit approach is not entirely convincing given the extensive financial resourcing that will be necessary to ensure weapons are predictable, reliable, and operationally compliant

⁷⁶ UK Ministry of Defence, Development, Concepts and Doctrine Centre, *The UK Approach to Unmanned Aircraft Systems*, Joint Doctrine Note 2/11 (30 March 2011) paras. 102-103.

⁷⁷ US Department of Defense, *Unmanned Systems Integrated Roadmap FY2011-2036* (2011) 44.

⁷⁸ US Department of Defense, *Unmanned Systems Integrated Roadmap FY2013-2038* (2013) 19 and 68.

⁷⁹ John Rawls, ‘Fifty Years after Hiroshima’ in Samuel Freeman (ed.), *Collected Papers: John Rawls* (Harvard 1999) 565-572.

with international humanitarian law.⁸⁰ There is also concern that a cost-benefit analysis is susceptible to efficiency-driven, short-term decisions which in the long-run may lead to combatant casualties and breaches of international humanitarian law.⁸¹

Closely related to the cost-benefit argument is the claim that due to the sophistication and superior capability of autonomous weapons in precision targeting, there are less likely to be civilian casualties and combatants will be taken out of harm's way. Apart from the speculative calculus, such a utilitarian rationale provides no interaction and interrelatedness between human combatants and human targets. It enables the former to treat the latter as irrational agents unworthy of human dignity. It is a one-sided justification made to the wider world without any relation to the individual target. What justification can be offered to the individual who is automatically treated as a lesser human? If there is no interaction and interrelatedness then there is potentially no restriction on the use of such means to conduct warfare. It may prove difficult to maintain the principle of distinction and categories of persons (for example hors de combat, POWs, civilians).⁸²

5. Conclusion

Assessing autonomous weapons against the Kantian notion of human dignity exposes their undesirability. First, ethical conduct of individuals and states should be determined by an inclusive moral theory; one providing justification for morals as well as legitimate responses to immoral conduct. Morality is rooted in something fundamental and objective: the human dignity of each person. Kant believed all humans, including wrongdoers, have rational capacity and autonomy of will which gives them an intrinsic value of dignity. This entitles humans to set and abide by rules, and to expect others to make rational choices. A rule that permits death by

⁸⁰ Alan Backstrom and Ian Henderson, 'New capabilities in warfare: an overview of contemporary technological developments and the associated legal and engineering issues in Article 36 weapons reviews' (2012) 94(886) *International Review of the Red Cross* 483-514, 508-509.

⁸¹ See, e.g., *Smith and Others v. Ministry of Defence* (2013) UKSC 41 negligence claims brought by military personnel against the UK Ministry of Defence for failures to provide target identity devices that allow automatic confirmation as to whether a vehicle is a friend or foe, and situational awareness equipment that permits tank crews to locate their position and direction of sight accurately.

⁸² Nagel argues that absolutist restrictions in warfare are based on a requirement that they be capable of specific justification to the person harmed rather than just to the world at large: Thomas Nagel, 'War and Massacre' (1972) 1 *Philosophy and Public Affairs* 128, 133-138.

autonomous weapons cannot be universalised because rational beings cannot freely and rationally agree that it respects human dignity. Autonomous weapons create a hierarchy of human dignity whereby human targets are unequal and unworthy of face-to-face human contact and, therefore, excluded from an inclusive moral theory. Their life is deemed less worthy than human combatants replaced by autonomous weapons, and they are treated as mere relative ends that can be subjected to efficient disposal.

Second, Kant's belief in not harming or killing others is not absolute and does allow for punishment with preconditions (for example judicial punishment and sentence, no mistreatment, limitations on methods and means of warfare). Human targets who have committed immoral acts maintain moral worth and human dignity so they cannot be mistreated or subject to punishment that dishonours humanity. Yet autonomous weapons do both. They treat humans as disposable inanimate objects and legitimise extrajudicial killing without giving the human an opportunity to avoid death. Although Geneva law prohibits inhumane treatment, lack of face-to-face killing and efficient disposal of human life make provisions against outrages upon the personal dignity of the dead redundant.

Rather than creating Kant's 'kingdom of ends' autonomous weapons perpetuate a cycle of irrationality and irrational agents; humans abdicate rational thinking on lethal force to a machine and human targets are violently ousted as irrational agents. It is a world of relative ends where there is no interaction and interrelatedness between human combatants and human targets. The perceived benefits of autonomous weapons are one-sided without any relation to the individual target. If there is no interaction and interrelatedness then there is potentially no restriction on the use of such means to conduct warfare, and it may prove difficult to maintain the principle of distinction and categories of persons.