

# SHAKING THE FOUNDATIONS

The Human Rights Implications of Killer Robots

HUMAN  
RIGHTS  
WATCH

**IHRC**

INTERNATIONAL HUMAN RIGHTS CLINIC

HUMAN RIGHTS PROGRAM AT HARVARD LAW SCHOOL





**IHRC**



# **Shaking the Foundations**

**The Human Rights Implications of Killer Robots**

Copyright © 2014 Human Rights Watch

All rights reserved.

Printed in the United States of America

ISBN: 978-1-62313-1333

Cover design by Rafael Jimenez

Human Rights Watch defends the rights of people worldwide. We scrupulously investigate abuses, expose the facts widely, and pressure those with power to respect rights and secure justice. Human Rights Watch is an independent, international organization that works as part of a vibrant movement to uphold human dignity and advance the cause of human rights for all.

Human Rights Watch is an international organization with staff in more than 40 countries, and offices in Amsterdam, Beirut, Berlin, Brussels, Chicago, Geneva, Goma, Johannesburg, London, Los Angeles, Moscow, Nairobi, New York, Paris, San Francisco, Sydney, Tokyo, Toronto, Tunis, Washington DC, and Zurich.

For more information, please visit our website: <http://www.hrw.org>

The International Human Rights Clinic (IHRC) at Harvard Law School seeks to protect and promote human rights and international humanitarian law through documentation; legal, factual, and strategic analysis; litigation before national, regional, and international bodies; treaty negotiations; and policy and advocacy initiatives. IHRC also critically examines the human rights movement and engages in innovative clinical education to develop advanced practice techniques and approaches to human rights advocacy. IHRC collaborates with leading international and local human rights organizations and bridges theory with practice at the law school while also advancing the interests of clients and affected communities around the world.

For more information, please visit IHRC's website: <http://hrp.law.harvard.edu/clinic/>



# Shaking the Foundations

## The Human Rights Implications of Killer Robots

**Summary ..... 1**

**Recommendations..... 4**

**I. State of the Debate..... 5**

**II. The Right to Life ..... 8**

    The Right to Life in Law Enforcement Situations..... 9

        Necessity ..... 10

        Exhaustion of All Alternatives ..... 11

        Proportionality ..... 12

    The Right to Life in Armed Conflict..... 14

**III. The Right to a Remedy.....17**

**IV. Human Dignity ..... 23**

**Conclusion ..... 25**

**Acknowledgments..... 27**



## Summary

Fully autonomous weapons, also called killer robots or lethal autonomous robots, do not yet exist, but weapons technology is moving rapidly toward greater autonomy. Fully autonomous weapons represent the step beyond remote-controlled armed drones. Unlike any existing weapons, these robots would identify and fire on targets without meaningful human intervention. They would therefore have the power to determine when to take human life.

The role of fully autonomous weapons in armed conflict and questions about their ability to comply with international humanitarian law have generated heated debate, but the weapons' likely use beyond the battlefield has been largely ignored. These weapons could easily be adapted for use in law enforcement operations, which would trigger international human rights law. This body of law generally has even more stringent rules than international humanitarian law, which is applicable only in situations of armed conflict.

Because human rights law applies during peace and war, it would be relevant to all circumstances in which fully autonomous weapons might be used. This report examines the weapons' human rights implications in order to ensure a comprehensive assessment of the benefits and dangers of fully autonomous weapons. The report finds that fully autonomous weapons threaten to violate the foundational rights to life and a remedy and to undermine the underlying principle of human dignity.

In 2013, the international community recognized the urgency of addressing fully autonomous weapons and initiated discussions in a number of forums. April marked the launch of the Campaign to Stop Killer Robots, an international civil society coalition coordinated by Human Rights Watch that calls for a preemptive ban on the development, production, and use of fully autonomous weapons. The following month, Christof Heyns, the special rapporteur on extrajudicial killing, submitted a report to the UN Human Rights Council that presented many objections to this emerging technology and called for national moratoria. In November, 117 states parties to the Convention on Conventional Weapons agreed to hold an experts meeting on what they refer to as lethal autonomous weapons systems in May 2014.

Human Rights Watch and Harvard Law School's International Human Rights Clinic (IHRC) have co-published a series of papers that highlight the concerns about fully autonomous weapons. In November 2012, they released *Losing Humanity: The Case against Killer Robots*, the first major civil society report on the topic, and they have since elaborated on the need for a new international treaty that bans the weapons.<sup>1</sup> While these earlier documents focus on the potential problems that fully autonomous weapons pose to civilians in war, this report seeks to expand the discussion by illuminating the concerns that use of the weapons in law enforcement operations raise under human rights law.

Fully autonomous weapons have the potential to contravene the right to life, which the Human Rights Committee describes as “the supreme right.”<sup>2</sup> According to the International Covenant on Civil and Political Rights (ICCPR), “No one shall be arbitrarily deprived of his life.”<sup>3</sup> Killing is only lawful if it meets three cumulative requirements for when and how much force may be used: it must be necessary to protect human life, constitute a last resort, and be applied in a manner proportionate to the threat. Each of these prerequisites for lawful force involves qualitative assessments of specific situations. Due to the infinite number of possible scenarios, robots could not be pre-programmed to handle every circumstance. In addition, fully autonomous weapons would be prone to carrying out arbitrary killings when encountering unforeseen situations. According to many roboticists, it is highly unlikely in the foreseeable future that robots could be developed to have certain human qualities, such as judgment and the ability to identify with humans, that facilitate compliance with the three criteria.

The use of fully autonomous weapons also threatens to violate the right to a remedy. International law mandates accountability in order to deter future unlawful acts and punish past ones, which in turn recognizes victims' suffering. It is uncertain, however, whether meaningful accountability for the actions of a fully autonomous weapon would be

---

<sup>1</sup> Human Rights Watch and Harvard Law School's International Human Rights Clinic (IHRC), *Losing Humanity: The Case against Killer Robots*, November 2012, <http://www.hrw.org/reports/2012/11/19/losing-humanity-o>; Human Rights Watch and IHRC, “Q&A on Fully Autonomous Weapons,” October 2013, <http://www.hrw.org/news/2013/10/21/qa-fully-autonomous-weapons>; Human Rights Watch and IHRC, “The Need for New Law to Ban Fully Autonomous Weapons,” November 2013, <http://www.hrw.org/news/2013/11/13/need-new-law-ban-fully-autonomous-weapons>.

<sup>2</sup> UN Human Rights Committee, General Comment 6, The Right to Life (Sixteenth session, 1982), Compilation of General Comments and General Recommendations Adopted by Human Rights Treaty Bodies, U.N. Doc. HRI/GEN/1/Rev.1 (1994), p. 6, para. 1. See also Manfred Nowak, *U.N. Covenant on Civil and Political Rights: CCPR Commentary* (Arlington, VA: N.P. Engel, 2005), p. 104.

<sup>3</sup> International Covenant on Civil and Political Rights (ICCPR), adopted December 16, 1966, G.A. Res. 2200A (XXI), 21 U.N. GAOR Supp. (No. 16) at 52, U.N. Doc. A/6316 (1966), 999 U.N.T.S. 171, entered into force March 23, 1976, art. 6(1).

possible. The weapon itself could not be punished or deterred because machines do not have the capacity to suffer. Unless a superior officer, programmer, or manufacturer deployed or created such a weapon with the clear intent to commit a crime, these people would probably not be held accountable for the robot's actions. The criminal law doctrine of superior responsibility, also called command responsibility, is ill suited to the case of fully autonomous weapons. Superior officers might be unable to foresee how an autonomous robot would act in a particular situation, and they could find it difficult to prevent and impossible to punish any unlawful conduct. Programmers and manufacturers would likely escape civil liability for the acts of their robots. At least in the United States, defense contractors are generally granted immunity for the design of weapons. In addition, victims with limited resources and inadequate access to the courts would face significant obstacles to bringing a civil suit.

Finally, fully autonomous weapons could undermine the principle of dignity, which implies that everyone has a worth deserving of respect. As inanimate machines, fully autonomous weapons could truly comprehend neither the value of individual life nor the significance of its loss. Allowing them to make determinations to take life away would thus conflict with the principle of dignity.

The human rights implications of fully autonomous weapons compound the many other concerns about use of the weapons. As Human Rights Watch and IHRC have detailed in other documents, the weapons would face difficulties in meeting the requirements of international humanitarian law, such as upholding the principles of distinction and proportionality, in situations of armed conflict. In addition, even if technological hurdles could be overcome in the future, failure to prohibit them now could lead to the deployment of models before their artificial intelligence was perfected and spark an international robotic arms race. Finally, many critics of fully autonomous weapons have expressed moral outrage at the prospect of humans ceding to machines control over decisions to use lethal force. In this context, the human rights concerns bolster the argument for an international ban on fully autonomous weapons.

## Recommendations

Based on the threats that fully autonomous weapons would pose to civilians during both law enforcement operations and armed conflict, Human Rights Watch and IHRC recommend that states:

- Prohibit the development, production, and use of fully autonomous weapons through an international legally binding instrument.
- Adopt national laws and policies that prohibit the development, production, and use of fully autonomous weapons.
- Take into account the human rights implications of fully autonomous weapons when discussing the weapons in international fora.

## I. State of the Debate

Fully autonomous weapons, also called killer robots or lethal autonomous robots, would revolutionize the “technology of killing.”<sup>4</sup> While traditional weapons are tools in the hand of a human being, fully autonomous weapons would make their own determinations about the use of lethal force. Unlike existing semi-autonomous armed drones, which are remotely controlled, fully autonomous weapons would have no human “in the loop.”

Fully autonomous weapon systems have yet to be created, but technology is moving rapidly in that direction. For example, the US X-47B prototype has taken off and landed on an aircraft carrier on its own, and a South Korean sentry robot can identify and shoot humans in the demilitarized zone with North Korea.<sup>5</sup> Neither is necessarily problematic at this stage—the former is not yet weaponized, and the latter requires a human to order it to fire—but such precursors demonstrate the evolution toward weapons with ever greater autonomy. Several other countries, including China, Israel, Russia, and the United Kingdom, have also allotted resources to developing this kind of technology.<sup>6</sup>

In 2013, the international community recognized the urgency of addressing fully autonomous weapons and initiated discussions in a number of forums. April marked the launch of the Campaign to Stop Killer Robots, an international civil society coalition coordinated by Human Rights Watch that calls for a preemptive ban.<sup>7</sup> The following month, Christof Heyns, the special rapporteur on extrajudicial killing, submitted a report to the UN Human Rights Council that presented many objections to this emerging technology; he called for national moratoria on production, transfer, acquisition, and use, and an independent panel to examine the issue more closely.<sup>8</sup> In November, 117 states parties to

---

<sup>4</sup> Eric Prokosch used the term when writing about the history of antipersonnel weapons. Eric Prokosch, *The Technology of Killing: A Military and Political History of Anti-Personnel Weapons* (London: Zed Books, 1995).

<sup>5</sup> Northrop Grumman, “X47B UCAS,” <http://www.northropgrumman.com/Capabilities/X47BUCAS/Pages/default.aspx> (accessed February 14, 2014); Brandon Vinson, “X-47B Makes First Arrested Landing at Sea,” *America’s Navy*, July 10, 2013, [http://www.navy.mil/submit/display.asp?story\\_id=75298](http://www.navy.mil/submit/display.asp?story_id=75298) (accessed February 16, 2014); Jon Rabirow, “Machine Gun-Toting Robots Deployed on DMZ,” *Stars and Stripes*, July 12, 2010, <http://www.stripes.com/news/pacific/korea/machine-gun-toting-robots-deployed-on-dmz-1.110809> (accessed February 14, 2014).

<sup>6</sup> Human Rights Watch and IHRC, *Losing Humanity*, pp. 7, 13-19.

<sup>7</sup> Campaign to Stop Killer Robots, “About Us,” <http://www.stopkillerrobots.org/about-us> (accessed February 14, 2014).

<sup>8</sup> UN Human Rights Council, Report of the Special Rapporteur on extrajudicial, summary or arbitrary executions, Christof Heyns, Lethal Autonomous Robotics, A/HRC/23/47, April 9, 2013,

the Convention on Conventional Weapons agreed to hold an informal experts meeting on what they refer to as lethal autonomous weapons systems in May 2014.<sup>9</sup>

Commentators have articulated strong but diverging views on fully autonomous weapons. Proponents maintain that the weapons would reduce the risk to soldiers' lives, decrease military expenditures, and be able to process information more quickly during operations. They also say that robots would be less apt to attack civilians because they would not act in fear or anger.<sup>10</sup> Opponents counter that fully autonomous weapons would likely endanger civilians. They argue that these weapons would lack compassion and empathy, important inhibitors to killing people needlessly, and they would not possess other human qualities, such as judgment, that are necessary to conduct the subjective assessments underlying many of international law's protections. In addition, opponents argue that it is unclear whether any person could in practice be held legally responsible for a robot's actions, and they assert that it is morally wrong to allow machines to make life-and-death determinations.<sup>11</sup>

Proponents of fully autonomous weapons contend that roboticists could theoretically develop technology with sensors to interpret complex situations and the ability to exercise near-human judgment.<sup>12</sup> Opponents question that assumption, emphasizing that such technology would not be possible in the foreseeable future, if ever. They also worry that deployment of such weapons could come well before the technological challenges were overcome and that allowing continued development would inevitably lead to a robotic arms race. Opponents believe that the existing legal, ethical, and scientific concerns raised by fully autonomous weapons outweigh speculation about the technology's potential benefits.<sup>13</sup>

---

[http://www.ohchr.org/Documents/HRBodies/HRCouncil/RegularSession/Session23/A-HRC-23-47\\_en.pdf](http://www.ohchr.org/Documents/HRBodies/HRCouncil/RegularSession/Session23/A-HRC-23-47_en.pdf) (accessed February 14, 2014).

<sup>9</sup> Meeting of the High Contracting Parties to the Convention on Conventional Weapons, Final Report, CCW/MSP/2013/10, November 15, 2013, para. 32.

<sup>10</sup> Ronald Arkin, *Governing Lethal Behavior in Autonomous Robots* (Boca Raton, FL: CRC Press, 2009), pp. 29-30.

<sup>11</sup> See, for example, Human Rights Watch and IHRC, *Losing Humanity*, pp. 30-45.

<sup>12</sup> Michael N. Schmitt and Jeffrey S. Thurnher, "'Out of the Loop': Autonomous Weapon Systems and the Law of Armed Conflict," *Harvard National Security Journal*, vol. 4 (2013), p. 239.

<sup>13</sup> Human Rights Watch and IHRC, "The Need for New Law to Ban Fully Autonomous Weapons," p. 12.

The majority of the debate so far has focused fully autonomous weapons in armed conflict, but once available, this technology could be adapted to a range of other contexts that can be grouped under the heading of law enforcement. For example, local police officers could potentially use such robots in crime fighting, the management of public protests, riot control, and other efforts to maintain law and order. State security forces could employ the weapons in attempts to control their opposition. Countries involved in international counter-terrorism could utilize them in scenarios that do not necessarily rise to the level of armed conflict as defined by international humanitarian law. Some law enforcement operations have legitimate ends, such as crime prevention; others, including violent suppression of peaceful protests, are inherently illegitimate. Fully autonomous weapons could be deployed in an operation regardless of its character.

The use of fully autonomous weapons in a law enforcement context would trigger the application of international human rights law. International human rights law applies in both peace and war, and it regulates the use of force in situations other than military operations and combat. In comparison to international humanitarian law, which governs military operations and combat and applies only during armed conflict, human rights law tends to have more stringent standards for regulating the use of lethal force, typically limiting it to where needed to defend human life and safety. Therefore, the challenges of developing a fully autonomous weapon that would comply with international law and still be useful are even greater when viewed through a human rights lens.

## II. The Right to Life

The right to life is the bedrock of international human rights law. The Universal Declaration of Human Rights (UDHR), the foundational document of this body of law, introduced the concept in 1948.<sup>14</sup> The International Covenant on Civil and Political Rights, a cornerstone human rights treaty, codified it in 1966.<sup>15</sup> Article 6 of the ICCPR states, “Every human being has the inherent right to life. This right shall be protected by law.”<sup>16</sup> Regional treaties from Africa, the Americas, and Europe also have incorporated the right to life.<sup>17</sup> In its General Comment 6, the Human Rights Committee, treaty body for the ICCPR, describes the right to life as “the supreme right” because it is a prerequisite for all other rights.<sup>18</sup> It is non-derogable even in public emergencies that threaten the existence of a nation.<sup>19</sup>

The right to life prohibits arbitrary killing. The ICCPR declares, “No one shall be arbitrarily deprived of his life.”<sup>20</sup> The Human Rights Committee states that this right should be interpreted broadly. ICCPR negotiators understood “arbitrary” as having legal and ethical meaning; for them, it encompassed unlawful and unjust acts.<sup>21</sup> While some drafters proposed enumerating permissible killings in the ICCPR, the group ultimately decided not to do so in order to emphasize the prohibition on arbitrary deprivation of life.<sup>22</sup>

---

<sup>14</sup> Universal Declaration of Human Rights (UDHR), adopted December 10, 1948, G.A. Res. 217A(III), U.N. Doc. A/810 at 71 (1948), art. 3 (“Everyone has the right to life, liberty and security of person.”).

<sup>15</sup> ICCPR, art. 6.

<sup>16</sup> *Ibid.*, art. 6(1).

<sup>17</sup> African [Banjul] Charter on Human and Peoples’ Rights, adopted June 27, 1981, OAU Doc. CAB/LEG/67/3 rev. 5, 21 I.L.M. 58 (1982), entered into force October 21, 1986, art. 4 (“Human beings are inviolable. Every human being shall be entitled to respect for his life and the integrity of his person. No one may be arbitrarily deprived of this right.”); American Convention on Human Rights (“Pact of San José, Costa Rica”), adopted November 22, 1969, O.A.S. Treaty Series No. 36, 1144 U.N.T.S. 123, entered into force July 18, 1978, reprinted in Basic Documents Pertaining to Human Rights in the Inter-American System, OEA/Ser.L.V/II.82 doc.6 rev.1 at 25 (1992), art. 4 (“Every person has the right to have his life respected. This right shall be protected by law and, in general, from the moment of conception. No one shall be arbitrarily deprived of his life.”); European Convention for the Protection of Human Rights and Fundamental Freedoms, 213 U.N.T.S. 222, entered into force September 3, 1953, as amended by Protocols Nos. 3, 5, 8, and 11 which entered into force on September 21, 1970, December 20, 1971, January 1, 1990, and November 1, 1998, respectively, art. 2 (“Everyone’s right to life shall be protected by law. No one shall be deprived of his life intentionally save in the execution of a sentence of a court following his conviction of a crime for which this penalty is provided by law.”).

<sup>18</sup> UN Human Rights Committee, General Comment 6, para. 1. See also Nowak, *U.N. Covenant on Civil and Political Rights*, p. 104.

<sup>19</sup> UN Human Rights Committee, General Comment 6, para. 1.

<sup>20</sup> ICCPR, art. 6(1).

<sup>21</sup> UN Human Rights Committee, General Comment 6, para. 1; Marc Bossuyt, ed., *Guide to the ‘travaux préparatoires’ of the International Covenant on Civil and Political Rights* (Dordrecht: Martinus Nijhoff Publishers, 1987), p. 122; Nowak, *U.N. Covenant on Civil and Political Rights*, pp. 110-111.

<sup>22</sup> Bossuyt, *Guide to the ‘travaux préparatoires’*, pp. 116-117.

## The Right to Life in Law Enforcement Situations

The right to life constrains the application of force in a range of situations outside of armed conflict. In its General Comment 6, the Human Rights Committee highlights the duty of states to prevent arbitrary killings by their security forces.<sup>23</sup> The United Nations set parameters for the use of force by such agents in the 1990 Basic Principles on the Use of Force and Firearms by Law Enforcement Officials (1990 Basic Principles) and the 1979 Code of Conduct for Law Enforcement Officials (1979 Code of Conduct). Adopted by a UN congress on crime prevention and the UN General Assembly respectively, these standards provide guidance for how to understand the scope of arbitrary killing in law enforcement situations.<sup>24</sup> They expressly note the importance of protecting human rights.<sup>25</sup>

Arbitrary killings under the right to life fail to meet three cumulative requirements for when and how much force may be used. To be lawful in law enforcement situations, force must be necessary, constitute a last resort, and be applied in a proportionate manner.<sup>26</sup> Fully autonomous weapons would face obstacles to meeting these criteria that circumscribe lawful force. They could not completely replicate the ability of human law enforcement officials to exercise judgment and compassion or to identify with other human beings, qualities that facilitate compliance with the law. These inherently human characteristics parallel those described in Article 1 of the UDHR, which states, “All human beings ... are endowed with reason and conscience and should act towards one another in a spirit of brotherhood.”<sup>27</sup> Due to the inadequacy of fully autonomous weapons in these areas, which is elaborated on below, the weapons could contravene the right to life, undermining the legitimacy of some law enforcement operations and exacerbating the harm caused by ones that are already illegitimate.

---

<sup>23</sup> UN Human Rights Committee, General Comment 6, para. 3.

<sup>24</sup> Basic Principles on the Use of Force and Firearms by Law Enforcement Officials, adopted by the Eighth United Nations Congress on the Prevention of Crime and the Treatment of Offenders (1990 Basic Principles), Havana, August 27 to September 7, 1990, U.N. Doc. A/CONF.144/28/Rev.1 at 112 (1990); United Nations Code of Conduct for Law Enforcement Officials (1979 Code of Conduct), adopted December 17, 1979, G.A. res. 34/169, annex, 34 U.N. GAOR Supp. (No. 46) at 186, U.N. Doc. A/34/46 (1979).

<sup>25</sup> 1990 Basic Principles, pmbll.; 1979 Code of Conduct, art. 2.

<sup>26</sup> The requirements for use of force can be broken down in different ways. For example, while covering almost the same the elements as this report does, the International Committee of the Red Cross (ICRC) divides the requirements into legality (force must be for a “legitimate” objective), necessity (there must be “no less restricting measures available”), proportionality (“any restriction of rights must be proportionate to the legitimate objective”), and precaution (“authorities must take all possible measures to minimize damage”). International Committee of the Red Cross, “Violence and the Use of Force,” July 2011, [http://www.icrc.org/eng/assets/files/other/icrc\\_002\\_0943.pdf](http://www.icrc.org/eng/assets/files/other/icrc_002_0943.pdf) (accessed February 16, 2013), p. 17.

<sup>27</sup> UDHR, art. 1.

## *Necessity*

Necessity is the first precondition for lawful force. The 1979 Code of Conduct states that law enforcement officials may employ force only when it is “strictly necessary” and “exceptional.”<sup>28</sup> The use of firearms is even more narrowly restricted to situations where it is essential to saving human lives. The 1990 Basic Principles limit officials’ use of firearms to:

self-defence or defence of others against the imminent threat of death or serious injury, to prevent the perpetration of a particularly serious crime involving grave threat to life, to arrest a person presenting such a danger or resisting their authority, or to prevent his or her escape.<sup>29</sup>

The 1990 Basic Principles add that “intentional lethal use of firearms may only be made when strictly unavoidable in order to protect life.”<sup>30</sup>

Fully autonomous weapons would lack human qualities that help law enforcement officials assess the seriousness of a threat and the need for a response. Both a machine and a police officer can take into account clearly visible signs, such as the presence of a weapon. However, interpreting more subtle cues whose meaning can vary by context, such as tone of voice, facial expressions, and body language, requires an understanding of human nature. A human officer can relate to another individual as a fellow human being, and this ability can help him or her read the other person’s intentions. Development of a fully autonomous weapon that could identify with the individual in the same way seems unlikely, and thus a robot might miss or misconstrue important clues as to whether a real threat to human life existed.<sup>31</sup>

---

<sup>28</sup> 1979 Code of Conduct, art. 3 and its commentary.

<sup>29</sup> 1990 Basic Principles, art. 9.

<sup>30</sup> Ibid. The 1979 Code of Conduct similarly states, “The use of firearms is considered an extreme measure.... In general, firearms should not be used except when a suspected offender offers armed resistance or otherwise jeopardizes the lives of others.” 1979 Code of Conduct, commentary to art. 3.

<sup>31</sup> See Marcello Guarini and Paul Bello, “Robotic Warfare: Some Challenges in Moving from Noncivilian to Civilian Theaters,” in *Robot Ethics: The Ethical and Social Implications of Robotics*, eds. Patrick Lin, Keith Abney, and George A. Bekey (Cambridge, MA: Massachusetts Institute of Technology, 2012), p. 138 (“A system without emotion ... could not predict the emotions or action of others based on its own states because it has no emotional states.”); Noel Sharkey, “Killing Made Easy: From Joysticks to Politics,” in *Robot Ethics: The Ethical and Social Implications of Robotics*, eds. Patrick Lin, Keith Abney, and George A. Bekey (Cambridge, MA: Massachusetts Institute of Technology, 2012), p. 118 (“Humans understand one another in a way that machines cannot. Cues can be very subtle, and there are an infinite number of circumstances where lethal force is inappropriate.”).

In addition, the deployment of fully autonomous weapons in law enforcement situations could affect the actions of the individual posing a potential threat. He or she might not know how to behave when confronted with a machine rather than a human law enforcement officer. The individual might respond differently to a robot than to a human and as a result unintentionally appear threatening. A robot's misinterpretation of the necessity of force could trigger an arbitrary killing in violation of the right to life.

### *Exhaustion of All Alternatives*

Law enforcement officials are required to exhaust all alternatives before applying force. The 1990 Basic Principles limit force and firearms to cases where “other means remain ineffective or without any promise of achieving the intended result.”<sup>32</sup> In particular, they permit the use of firearms exclusively “when less extreme means are insufficient” to save a life.<sup>33</sup> These standards encourage states to develop non-lethal weapons and equip law enforcement officials with self-defense equipment in order to decrease the need to employ lethal force.<sup>34</sup>

Fully autonomous weapons' inability to relate to humans could interfere with their ability to ensure that all means short of force are exhausted. The 1990 Basic Principles state that training law enforcement officials in “methods of persuasion, negotiation, and mediation” is important to decreasing the use of lethal force.<sup>35</sup> To deescalate a situation, human officers often appeal to a threatening individual's reason, emotions, and interests. At present there is little prospect of developing a fully autonomous weapon that would be “intelligent” enough to be able to “talk down” an individual and defuse a standoff. A potential perpetrator would be more apt to connect with and be persuaded by a fellow human than an inanimate machine. Furthermore, it is unlikely that a fully autonomous weapon would be able to read a situation well enough to strategize about the best alternatives to use of force. On a more practical level, even if fully autonomous systems could be equipped with non-lethal weapons, the possibility of developing robots with the capability to restrain individuals or take prisoners seems remote. Fully autonomous

---

<sup>32</sup> 1990 Basic Principles, art. 4.

<sup>33</sup> Ibid., art. 9. The 1979 Code of Conduct says that firearms may only be used when “less extreme measures are not sufficient.” 1979 Code of Conduct, commentary to art. 3.

<sup>34</sup> 1990 Basic Principles, art. 2.

<sup>35</sup> Ibid., art. 20.

weapons could thus escalate a situation before demonstrating that methods short of force are “ineffective or without any promise of achieving the intended result.”<sup>36</sup>

### *Proportionality*

International law enforcement standards also specify how much force may be used when it is necessary and all other means have been exhausted. Force must be proportional to the threat involved. The 1990 Basic Principles require law enforcement officials to “act in proportion to the seriousness of the offence and the legitimate objective to be achieved.”<sup>37</sup> They oblige officials to “exercise restraint” and to minimize the harm they cause.<sup>38</sup> For example, in dispersing violent assemblies, officials may use lethal weapons only “to the minimum extent necessary.”<sup>39</sup> The 1979 Code of Conduct similarly highlights the principle of proportionality and states that law enforcement officials’ use of force must be “to the extent required for the performance of their duty.”<sup>40</sup>

Choosing an appropriate level of force could pose additional problems for fully autonomous weapons. First, these weapons would not possess judgment, which human officers rely on to balance the force of the response with the gravity of the perceived threat. The Oxford English Dictionary defines judgment as “the ability to make considered decisions or to arrive at reasonable conclusions or opinions on the basis of the available information.”<sup>41</sup> Judgment requires human capabilities of reason and reflection to interpret information and formulate an opinion. In law enforcement, judgment allows officers to assess complicated situations, taking into account such factors as a perpetrator’s background, mental state, and demands, and then to make case-by-case decisions about the minimum level of force necessary.

There are serious doubts that fully autonomous weapons could determine how much force is proportionate in a particular case. A designer could not pre-program a robot to deal with all situations because even a human could not predict the infinite possibilities. In addition,

---

<sup>36</sup> *Ibid.*, art. 4.

<sup>37</sup> *Ibid.*, art. 5(a).

<sup>38</sup> *Ibid.*, art. 5(a) and (b).

<sup>39</sup> *Ibid.*, art. 14.

<sup>40</sup> 1979 Code of Conduct, art. 3 and its commentary.

<sup>41</sup> Oxford English Dictionary online, “Judgment.” Black’s Law Dictionary defines judgment as “[t]he formation of an opinion or notion concerning some thing by exercising the mind upon it.” *Black’s Law Dictionary*, abridged sixth ed., “Judgment” (St. Paul, MN: West Publishing Co., 1991).

it would be difficult for a fully autonomous weapon to replicate the complex, subjective thinking processes required to judge unforeseen circumstances. Proportionality determinations involve more than quantitative analysis, and according to Christof Heyns, the UN special rapporteur on extrajudicial killing, “While robots are especially effective at dealing with quantitative issues, they have limited abilities to make the qualitative assessments that are often called for when dealing with human life.”<sup>42</sup>

Second, fully autonomous weapons would lack emotions that can generate the kind of restraint the 1990 Basic Principles oblige law enforcement officials to exercise. While fully autonomous weapons would not respond to threats in fear or anger, they would also not feel the “natural inhibition of humans not to kill or hurt fellow human beings.”<sup>43</sup> Studies of human soldiers have demonstrated that “there is within man an intense resistance to killing their fellow man.”<sup>44</sup> Compassion contributes to such a resistance, but it is hard to see how the capacity to feel compassion could be reproduced in robots. Human rights law does not require the exercise of compassion in particular, but the emotion can facilitate compliance with proportionality and serve as a safeguard against the disproportionate use of force.<sup>45</sup>

Abusive autocrats could take advantage of fully autonomous weapons’ lack of inherent restraint. For example, they could deploy these weapons to suppress protestors with a level of violence against which human security forces might rebel. Even the most hardened troops can eventually turn on their leader if ordered to fire on their own people. An abusive leader who resorted to fully autonomous weapons would be free of the fear that security forces would resist being deployed against certain targets.

---

<sup>42</sup> UN Human Rights Council, Report of the Special Rapporteur on extrajudicial, summary or arbitrary executions, Christof Heyns, *Lethal Autonomous Robotics*, pp. 10-11.

<sup>43</sup> Armin Krishnan, *Killer Robots: Legality and Ethicality of Autonomous Weapons* (Farnham, UK: Ashgate Publishing Limited, 2009), p. 130.

<sup>44</sup> Lt. Col. Dave Grossman, *On Killing: The Psychological Cost of Learning to Kill in War and Society* (New York: Little, Brown and Company, 1995), p. 4. For example, based on interviews with thousands of US soldiers in World War II, US Army Brig. Gen. S.L.A. Marshall found that usually only 15 to 20 percent of troops would fire at the enemy. These numbers were due to an innate hesitancy to kill, not to fear or cowardice because “[t]hose who would not fire did not run or hide (and in many cases they were willing to risk great danger to rescue comrades, get ammunition, or run messages).” S.L.A. Marshall, *Men Against Fire: The Problem of Battle Command in Future War* (New York: William Morrow & Company, 1947), p. 54; Grossman, *On Killing*, p. 4. See also Grossman, *On Killing*, pp. 16-17 (discussing a 1986 study by the British Defense Operational Analysis Establishment of 100 “nineteenth- and twentieth-century battles and test trials”).

<sup>45</sup> For a discussion of the dangers posed by fully autonomous weapons’ lack of compassion, see Human Rights Watch and IHRC,  *Losing Humanity*, pp. 37-39.

It is highly unlikely therefore that fully autonomous weapons would be able to comply with the three requirements for lawful use of force—necessity, exhaustion of all alternatives, and proportionality. As a result, this technology would have the potential to arbitrarily deprive innocent people of their lives in law enforcement situations.

## The Right to Life in Armed Conflict

Because the right to life is non-derogable even in situations that threaten a country's existence, the right continues to apply during armed conflict.<sup>46</sup> Under circumstances of armed conflict, many look to international humanitarian law, the law that governs that specific context (*lex specialis*), to help interpret certain human rights provisions.<sup>47</sup> Therefore, in wartime, arbitrary killing refers to unlawful killing under international humanitarian law. In his authoritative commentary on the ICCPR, Manfred Nowak, former UN special rapporteur on torture, defines arbitrary killings in armed conflict as “those that contradict the humanitarian laws of war.”<sup>48</sup> The International Committee of the Red Cross (ICRC) Customary International Humanitarian Law Database states, “The prohibition of ‘arbitrary deprivation of the right to life’ under human rights law ... encompasses unlawful killing in the conduct of hostilities.”<sup>49</sup> The ICRC finds unlawful killings go beyond those

---

<sup>46</sup> In its advisory opinion on nuclear weapons, the International Court of Justice (ICJ) found, “In principle, the right not arbitrarily to be deprived of one’s life applies also in hostilities.” Legality of the Threat or Use of Nuclear Weapons, ICJ, Advisory Opinion, July 8, 1996, para. 25 [hereinafter Nuclear Weapons Advisory Opinion]. See also Nowak, *U.N. Covenant on Civil and Political Rights*, p. 108 (“Arbitrary killings in the course of armed conflicts permissible under international law and civil wars also represent a violation of the right to life.”); UN Human Rights Committee, General Comment 31, The Nature of the General Legal Obligation Imposed on States Parties to the Covenant (Eightieth session, 2004), U.N. Doc. CCPR/C/21/Rev.1/Add.13 (2004), para. 11 (The ICCPR “applies also in situations of armed conflict to which the rules of international humanitarian law are applicable.”).

<sup>47</sup> For further discussion of the debate about the relationship between international human rights law and international humanitarian law, see generally Noëlle Quénivet, “The History of the Relationship Between International Humanitarian Law and Human Rights Law,” in *International Humanitarian Law and Human Rights Law: Towards a New Merger in International Law*, eds. Roberta Arnold and Noëlle Quénivet (Leiden: Martinus Nijhoff, 2008). According to the ICJ, “The test of what is an arbitrary deprivation of life ... falls to be determined by the applicable *lex specialis*, namely, the law applicable in armed conflict which is designed to regulate the conduct of hostilities.” Nuclear Weapons Advisory Opinion, para. 25. Recognizing a link between the two bodies of law, the Human Rights Committee stated in General Comment 31 that “in respect of certain Covenant rights, more specific rules of international humanitarian law may be specially relevant for the purposes of the interpretation of Covenant rights,” though it noted that “both spheres of law are complementary, not mutually exclusive.” UN Human Rights Committee, General Comment 31, para. 11.

<sup>48</sup> Nowak, *U.N. Covenant on Civil and Political Rights*, p. 108, n. 29. Given that states “have the supreme duty to prevent wars,” killings in the course of a war that violates the UN Charter would also violate the right to life. *Ibid.*, p. 108; UN Human Rights Committee, General Comment 6, para. 2.

<sup>49</sup> ICRC, Customary International Humanitarian Law Database, “Rule 89: Violence to Life,” [http://www.icrc.org/customary-ihl/eng/docs/v1\\_rul\\_rule89](http://www.icrc.org/customary-ihl/eng/docs/v1_rul_rule89), (accessed February 14, 2014).

violations that are considered grave breaches or war crimes, such as direct attacks against civilians, to cover indiscriminate and disproportionate attacks.<sup>50</sup>

Civilian protection in international humanitarian law rests on the rules of distinction and proportionality. The former requires parties to a conflict to distinguish between combatants and civilians, and it outlaws means or methods of war that “cannot be directed at a specific military objective” because they are indiscriminate.<sup>51</sup> Proportionality, a subset of distinction, prohibits attacks in which expected civilian harm “would be excessive” compared to the anticipated military advantage.<sup>52</sup> “Proportionality” has a somewhat different, although not contradictory, meaning in international humanitarian law than it does in international human rights law. While under human rights law the term regulates the level of force acceptable to respond to a threat, under international humanitarian law it is used to judge whether an offensive or defensive military action is lawful. Both rules aim to protect human lives.

Some of the limitations of fully autonomous weapons that raise concerns in the law enforcement context, such as the inability to identify with humans or exercise human judgment, could also interfere with the weapons’ compliance with international law during armed conflict. First, there are serious doubts about whether fully autonomous weapons could distinguish adequately between combatants and civilians. Enemy combatants in contemporary conflicts often shed visible signs of military status, such as uniforms, which makes recognizing their intentions crucial to differentiating them from civilians.<sup>53</sup> As discussed above, it would be challenging to program fully autonomous weapons to understand human intentions. Because as machines they could not identify with humans,

---

<sup>50</sup> “[U]nlawful killings can result, for example, from a direct attack against a civilian (see Rule 1), from an indiscriminate attack (see Rule 11) or from an attack against military objectives causing excessive loss of civilian life (see Rule 14), all of which are prohibited by the rules on the conduct of hostilities.” Ibid.

<sup>51</sup> Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of International Armed Conflicts (Protocol I), adopted June 8, 1977, 1125 U.N.T.S. 3, entered into force December 7, 1978, arts. 48 and 51(4)(b). The protocol also prohibits indiscriminate attacks that “are not directed at a specific military objective” or “employ a method or means of combat the effects of which cannot be limited as required by this Protocol; and consequently, in each such case, are of a nature to strike military objectives and civilians or civilian objects without distinction.” Ibid., art. 51(4)(a) and (c).

<sup>52</sup> Ibid., art. 51(5)(b)

<sup>53</sup> Guarini and Bello, “Robotic Warfare,” p. 131 (“In a context where we cannot assume that everyone present is a combatant, then we have to figure out who is a combatant and who is not. This frequently requires the attribution of intention.”).

they would find it more difficult to recognize and interpret subtle behavioral clues whose meaning depends on context and culture.<sup>54</sup>

Second, fully autonomous weapons could face obstacles in making targeting choices that accord with the proportionality test. These weapons could not be pre-programmed to deal with every type of situation, and deciding how to weigh civilian harm and military advantage in unanticipated situations “involve[s] distinctively human judgement [sic],” something it is doubtful machines could replicate.<sup>55</sup>

Although the ability of fully autonomous weapons to process complex information might improve in the future, it seems implausible that they could ever be identical to humans. As a result, these weapons would find it difficult to meet the three criteria for use of force in law enforcement or comply with the rules of distinction and proportionality in armed conflict. Fully autonomous weapons would have the potential to kill arbitrarily and thus violate the right that underlies all others, the right to life.

---

<sup>54</sup> Human Rights Watch and IHRC, *Losing Humanity*, pp. 30-32; Guarini and Bello, “Robotic Warfare,” p. 138; Sharkey, “Killing Made Easy,” p. 118.

<sup>55</sup> UN Human Rights Council, Report of the Special Rapporteur on extrajudicial, summary or arbitrary executions, Christof Heyns, *Lethal Autonomous Robotics*, pp. 32-34 (noting that because the proportionality test is a subjective one, it requires human judgment, “rather than the automatic decision making characteristics of a computer.”). In addition, proportionality decisions are generally assessed according to the reasonable military commander standard. According to the Max Planck Encyclopedia of International Law, “[t]he concept of reasonableness exhibits an important link with *human* reason,” and is “generally perceived as opening the door to several *ethical or moral*, rather than legal, considerations.” Many people would object to the idea that machines can or should be making ethical or moral calculations, and yet this is precisely what the reasonable military commander standard requires. Max Planck Encyclopedia of Public International Law online, “Reasonableness in International Law” (emphasis added).

### III. The Right to a Remedy

The right to a remedy applies to violations of all human rights. The UDHR lays out the right, and Article 2(3) of the ICCPR requires states parties to “ensure that any person whose rights or freedoms ... are violated shall have an effective remedy.”<sup>56</sup> Several regional human rights treaties incorporate it as well.<sup>57</sup>

The right to a remedy requires states to ensure individual accountability. It includes the duty to prosecute individuals for serious violations of human rights law. In its General Comment 31, the Human Rights Committee explains that the ICCPR obliges states parties to investigate allegations of wrongdoing and, if they find evidence of certain types of violations, to bring perpetrators to justice.<sup>58</sup> A failure to investigate and, where appropriate, prosecute “could in and of itself give rise to a separate breach of the Covenant.”<sup>59</sup> The 2005 Basic Principles and Guidelines on the Right to a Remedy and Reparation (2005 Basic Principles and Guidelines), standards adopted by the UN General Assembly, reiterate the obligation to investigate and prosecute. They also require states to punish individuals who are found guilty.<sup>60</sup>

The duty to prosecute applies to acts committed in law enforcement situations or armed conflict. The 2005 Basic Principles and Guidelines require states to prosecute gross violations of international human rights law, and the Human Rights Committee includes arbitrary killings among those crimes.<sup>61</sup> The 2005 Basic Principles and Guidelines also

---

<sup>56</sup> UDHR, art. 8 (“Everyone has the right to an effective remedy by the competent national tribunals for acts violating the fundamental rights granted him by the constitution or by law.”); ICCPR, art. 2(3).

<sup>57</sup> African Charter on Human and People’s Rights, art. 7(1) (“Every individual shall have the right to have his cause heard. This comprises: (a) the right to an appeal to competent national organs against acts of violating his fundamental rights as recognized and guaranteed by conventions, laws, regulations and customs in force.”); American Convention on Human Rights, art. 25 (“Everyone has the right to simple and prompt recourse, or any other effective recourse, to a competent court or tribunal for protection against acts that violate his fundamental rights recognized by the constitution or laws of the state concerned or by this Convention.”); European Convention for the Protection of Human Rights and Fundamental Freedoms, art. 13 (“Everyone whose rights and freedoms as set forth in this Convention are violated shall have an effective remedy before a national authority.”).

<sup>58</sup> UN Human Rights Committee, General Comment 31, paras. 15 and 18.

<sup>59</sup> *Ibid.*

<sup>60</sup> Basic Principles and Guidelines on the Right to a Remedy and Reparation for Victims of Gross Violations of International Human Rights Law and Serious Violations of International Humanitarian Law (2005 Basic Principles and Guidelines), UN General Assembly, Resolution 60/47, December 16, 2005, art. 4.

<sup>61</sup> *Ibid.*, art. 4; Human Rights Committee, General Comment 31, para. 18.

cover “serious violations of international humanitarian law,” the *lex specialis* for armed conflict.<sup>62</sup> The Fourth Geneva Convention and its Additional Protocol I, international humanitarian law’s key civilian protection instruments, similarly oblige states to prosecute “grave breaches,” i.e., war crimes, such as intentionally targeting civilians or knowingly launching a disproportionate attack.<sup>63</sup>

The right to a remedy is not limited to criminal prosecution. It encompasses reparations, which can include “restitution, compensation, rehabilitation, satisfaction and guarantees of non-repetition.”<sup>64</sup> States have the responsibility to provide many of these reparations. The 2005 Basic Principles and Guidelines, however, also oblige states to enforce judgments related to claims brought by victims against individuals or entities.<sup>65</sup> These standards “are without prejudice to the right to a remedy and reparation” for *all* violations of international human rights and humanitarian law, not just crimes.<sup>66</sup> Victims are thus entitled to some form of remedy, including but not limited to criminal prosecution.

Accountability serves a dual policy purpose. First, it seeks to deter future violations of the law. According to the Human Rights Committee, “the purposes of the Covenant would be defeated without an obligation ... to take measures to prevent a recurrence of a violation.”<sup>67</sup> Second, a remedy serves as retribution, which provides victims the satisfaction that someone was punished for the harm they suffered. Dinah Shelton, author of an influential commentary on remedies, wrote that “punishment conveys to criminals and others that they wronged the victim and thus implicitly recognizes the victims [sic] plight and honors the victim’s moral claims.”<sup>68</sup> Applying these principles to the context of fully autonomous weapons, Christof Heyns, the special rapporteur on extrajudicial killing, declared that if responsibility for a robot’s actions is impossible, “its use should be considered unethical and unlawful as an abhorrent weapon.”<sup>69</sup>

---

<sup>62</sup> 2005 Basic Principles and Guidelines, art. 4.

<sup>63</sup> Geneva Convention relative to the Protection of Civilian Persons in Time of War (Fourth Geneva Convention), adopted August 12, 1949, 75 U.N.T.S. 287, entered into force October 21, 1950, art. 146; Protocol I, art. 86.

<sup>64</sup> 2005 Basic Principles and Guidelines, arts. 18-23.

<sup>65</sup> *Ibid.*, arts. 17 and 15.

<sup>66</sup> *Ibid.*, art. 26

<sup>67</sup> Human Rights Committee, General Comment 31, para. 17. See also Dinah Shelton, *Remedies in International Human Rights Law* (Oxford: Oxford University Press, 2005), pp. 13-14.

<sup>68</sup> Shelton, *Remedies in International Human Rights Law*, p. 13.

<sup>69</sup> UN Human Rights Council, Report of the Special Rapporteur on extrajudicial, summary or arbitrary executions, Christof Heyns, *Lethal Autonomous Robotics*, p. 15.

In both law enforcement operations and armed conflict, the actions of fully autonomous weapons would likely fall within an accountability gap that would contravene the right to a remedy. It is unclear who would be liable when an autonomous machine makes life-and-death determinations about the use of force without meaningful human intervention. Assigning responsibility to the robot would make little sense. It could not be punished as a human can because it could not experience physical or psychological pain.<sup>79</sup> As discussed below, significant legal and practical obstacles exist to holding accountable the other most likely candidates—the superior officer or commander, programmer, and manufacturer—not the least because it will be difficult to foresee all situations the robot might face, and thus predict and account for its actions. In these cases, it is hard to see how one could achieve accountability under existing law in a way that was not unfair to any human who might be accused. The accountability gap that resulted would likely leave victims frustrated that no one was held accountable and punished for their suffering. It would also undermine the policy goals of deterrence and retribution.

If a superior officer, a programmer, or a manufacturer used or created a fully autonomous weapon with the clear intent to violate the right to life, they could most likely be held directly criminally liable. Criminal intent would be difficult to prove, however, and presumably be rare, at least among representatives of a state that generally abides by international law. Of greater concern would be a situation in which a fully autonomous weapon committed an arbitrary killing but there was no evidence a human intended or foresaw it. In such a case, there would be no human to hold directly responsible for the decision to attack, and indirect liability would be difficult to achieve.

Superior officers or commanders are generally not considered accountable for the actions of their subordinates because the latter make autonomous choices, as fully autonomous weapons would. The principle of superior responsibility, also known as command responsibility, is the primary exception to this rule. According to the 1990 Basic Principles:

Governments and law enforcement agencies shall ensure that superior officers are held responsible if they know, or should have known, that law enforcement officials under their

---

<sup>79</sup> Robert Sparrow, “Killer Robots,” *Journal of Applied Philosophy*, vol. 24 (2007), p. 72.

command are resorting, or have resorted, to the unlawful use of force and firearms, and they did not take all measures in their power to prevent, suppress or report such use.<sup>71</sup>

Command responsibility similarly holds military commanders responsible for subordinates' actions if they knew or should have known their subordinates committed or were going to commit a crime and failed to prevent the crime or punish the subordinates.<sup>72</sup>

The doctrine of superior responsibility is ill suited to establishing accountability for the actions of a fully autonomous weapon. First, it would be difficult for a superior officer to possess the requisite knowledge of an autonomous robot's future actions. These actions would often be unforeseeable, especially given that international human rights law and international humanitarian law frequently require context-specific determinations and that in many cases superior officers might not have the scientific expertise to predict how such a complex piece of technology would function. Second, a superior officer who deployed a fully autonomous weapon would face obstacles to preventing or punishing a robot's unlawful actions. The superior officer would be unable to prevent them if he or she could not foresee how the machine might act in different situations. The superior officer could not punish a fully autonomous weapon after the fact since, as discussed above, a robot cannot feel pain. Unless all of the elements of superior responsibility were met, a superior officer could not be held legally responsible for the actions of a fully autonomous weapon.

Alternatively, the law could try to hold a programmer or manufacturer responsible for the acts of a fully autonomous weapon. Civil tort law offers an approach other than prosecution, but it too would likely fail to ensure the right to a remedy. In the United States, for example, defense contractors are generally not found liable for harm caused by their products. Under the Federal Torts Claims Act, the government waives its immunity from civil suits in certain situations. The Supreme Court has applied this rule to contractors hired by the government. The waiver, however, is subject to the discretionary function exception and the combatant activities exception.<sup>73</sup> The first grants immunity for design defects in military equipment when:

---

<sup>71</sup> 1990 Basic Principles, art. 24.

<sup>72</sup> See, for example, Rome Statute of the International Criminal Court (Rome Statute), U.N. Doc. A/CONF.183/9, July 17, 1998, entered into force July 1, 2002, art. 28; Statute of the International Criminal Tribunal for the former Yugoslavia (ICTY Statute), S.C. Res. 827, U.N. Doc. S/RES/827 (1993), as amended, art. 7(3).

<sup>73</sup> "There are two important exceptions at the heart of the military contractor's efforts to invoke federal sovereign immunity. The first exception is the 'discretionary-function' exception, which preserves federal sovereign immunity in suits arising from

(a) the United States approved reasonably precise specifications; (b) the equipment conformed to those specifications; and (c) the supplier warned the United States about dangers in the use of the equipment that were known to the supplier but not to the United States.<sup>74</sup>

The second, the combat activities exception, states that contractors have “no duty of reasonable care ... to those against whom force is directed as a result of authorized military action” and that litigation should not lead to the disclosure of secret weapon designs.<sup>75</sup> The programming and manufacturing of a fully autonomous weapon could fall under at least one of these exceptions, allowing the robot’s creators to escape liability. These two exceptions apply only in the United States, but they are significant because the United States is a leader in the development of autonomous weapons technology. Like the limits of superior responsibility, immunity under tort law could present an obstacle to holding individuals accountable for the actions of fully autonomous weapons.

Even without a legal gap, there are policy and practical problems with holding programmers and manufacturers accountable. Such liability could be unfair since even programmers and manufacturers might be unable to foresee the harm their fully autonomous weapons could cause in various situations. One commentator wrote that assigning them responsibility would be like “holding parents accountable for the actions of their children after they have left their care.”<sup>76</sup> Liability could also create a moral hazard whereby superior officers become more likely to deploy the weapons systems in dangerous situations because they believe programmers and manufacturers would bear any responsibility. Finally, civil suits are generally brought by victims, and it is unrealistic

---

‘the exercise or performance or the failure to exercise or perform a discretionary function’ by the federal government or its employees.... The second exception is the ‘combatant-activities’ exception, which excepts suits against the federal government ‘arising out of the combatant activities of the military or naval forces, or the Coast Guard, during time of war.’” Andrew Finkelman, “Suing the Hired Guns: An Analysis of Two Federal Defenses to Tort Lawsuits against Military Contractors,” *Brooklyn Journal of International Law*, vol. 34 (2009), p. 405.

<sup>74</sup> *Boyle v. United Technologies Corp.*, 487 U.S. 500, 511 (1988).

<sup>75</sup> *Koohi v. United States*, 976 F.2d 1328, 1337 (9th Cir. 1992); *Bentzlin v. Hughes Aircraft Co.*, 833 F. Supp. 1486, 1490 (C.D. Cal. 1993). In *Koohi*, the court found that “no duty of reasonable care” was owed to civilian victims of the accidental downing of an Iranian civilian aircraft. In *Bentzlin*, the court found manufacturer immunity in a case of six friendly fire deaths. Courts distinguished some later cases from *Koohi* and *Bentzlin* based on factual differences but would likely grant immunity in cases involving fully autonomous weapons used against enemy forces given the national security interests in protecting the robots’ complex designs. See, for example, *Carmichael v. Kellogg, Brown & Root Servs., Inc.*, 450 F. Supp. 2d 1373, 1379-1380 (N.D. Ga. 2006).

<sup>76</sup> Sparrow, “Killer Robots,” *Journal of Applied Philosophy*, p. 70. See generally Andreas Matthias, “The Responsibility Gap: Ascribing Responsibility for the Actions of Learning Automata,” *Ethics and Information Technology*, vol. 6 (2004), p. 175-183.

to think all victims would have the resources or adequate access to obtain justice.<sup>77</sup> This practical limitation is significant because civil litigation against those who program, manufacture, or use such robots would be a more likely avenue of redress than criminal prosecution. Lack of accountability for superior officers, programmers, and manufacturers would thus interfere with victims exercising their right to a remedy.

---

<sup>77</sup> UN Human Rights Committee, Report of the Special Rapporteur on extrajudicial, summary or arbitrary executions, Christof Heyns, *Lethal Autonomous Robotics*, p. 15.

## IV. Human Dignity

The concept of human dignity lies at the heart of international human rights law. The opening words of the UDHR assert that “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.”<sup>78</sup> In ascribing inherent dignity to all human beings, the UDHR implies that everyone has worth that deserves respect.<sup>79</sup> The ICCPR establishes the inextricable link between dignity and human rights, stating in its preamble that the rights it enumerates “derive from the inherent dignity of the human person.”<sup>80</sup> Regional treaties echo this position, and the Vienna Declaration of the 1993 World Human Rights Conference affirms that “all human rights derive from the dignity and worth inherent in the human person.”<sup>81</sup>

Fully autonomous weapons would possess the power to kill people yet be unable to respect their dignity. As inanimate machines, they could truly comprehend neither the value of individual life nor the significance of its loss. Allowing them to make determinations to take life away would thus conflict with the principle of dignity.<sup>82</sup>

Critics of fully autonomous weapons have expressed serious moral concerns related to these shortcomings. In his 2013 report to the Human Rights Council, Christof Heyns, the special rapporteur on extrajudicial killing, wrote,

---

<sup>78</sup> UDHR, pmb., para. 1 (emphasis added). The Oxford English Dictionary defines dignity as “the quality of being worthy or honourable; worthiness, worth, nobleness, excellence.” Oxford English Dictionary online, “Dignity.”

<sup>79</sup> Jack Donnelly, “Human Dignity and Human Rights,” in Swiss Initiative to Commemorate the 60<sup>th</sup> Anniversary of the UDHR, *Protecting Dignity: Agenda for Human Rights*, June 2009, [http://www.udhr60.ch/report/donnelly-HumanDignity\\_0609.pdf](http://www.udhr60.ch/report/donnelly-HumanDignity_0609.pdf) (accessed February 16, 2014), p. 10.

<sup>80</sup> ICCPR, pmb., para. 2.

<sup>81</sup> African Charter on Human and People’s Rights, pmb., para. 2, and art. 5 (noting that dignity is one of the “essential objectives for the achievement of the legitimate aspirations of the African peoples” and granting the right to respect of dignity); American Convention on Human Rights, art. 5 (referring to the “inherent dignity of the human person”); European Convention for the Protection of Human Rights and Fundamental Freedoms, pmb., para 1 (declaring that the right to life is “essential ... for the full recognition of the inherent dignity of all human beings”); Vienna Declaration and Program of Action, adopted by the World Conference on Human Rights, June 25, 1993, pmb., para. 2.

<sup>82</sup> According to UN Special Rapporteur Christof Heyns, “[T]here is widespread concern that allowing [fully autonomous weapons] to kill people may denigrate the value of life itself.” UN Human Rights Council, Report of the Special Rapporteur on extrajudicial, summary or arbitrary executions, Christof Heyns, Lethal Autonomous Robotics, p. 20.

[A] human being somewhere has to take the decision to initiate lethal force and as a result internalize (or assume responsibility for) the cost of each life lost in hostilities, as part of a deliberative process of human interaction.... Delegating this process dehumanizes armed conflict even further and precludes a moment of deliberation in those cases where it may be feasible. Machines lack morality and mortality, and should as a result not have life and death powers over humans.<sup>83</sup>

Heyns described this issue as an “overriding consideration” and declared that if fully autonomous weapons are found morally unacceptable, “no other consideration can justify the deployment of [fully autonomous weapons], no matter the level of technical competence at which they operate.”<sup>84</sup>

---

<sup>83</sup> Ibid., p. 17.

<sup>84</sup> Ibid.

## Conclusion

Fully autonomous weapons threaten to contravene foundational elements of human rights law. They could violate the right to life, a prerequisite for all other rights. Deficiencies in judgment, compassion, and capacity to identify with human beings could lead to arbitrary killing of civilians during law enforcement or armed conflict operations. Fully autonomous weapons could also cause harm for which individuals could not be held accountable, thus undermining the right to a remedy. Robots could not be punished, and superior officers, programmers, and manufacturers would all be likely to escape liability. Finally, as machines, fully autonomous weapons could not comprehend or respect the inherent dignity of human beings. The inability to uphold this underlying principle of human rights raises serious moral questions about the prospect of allowing a robot to take a human life.

Proponents of fully autonomous weapons might argue that technology could eventually help address the problems identified in this report, and it is impossible to know where science will lead.<sup>85</sup> In a 2013 public letter, however, more than 270 roboticists, artificial intelligence experts, and other scientists expressed their skepticism that adequate developments would be possible.<sup>86</sup> Given this uncertainty, the potential of fully autonomous weapons to violate human rights law, combined with other ethical, legal, policy, and scientific concerns, demands a precautionary approach. The precautionary principle of international law states that “[w]here there are threats of serious or irreversible damage, lack of full scientific certainty shall not be used as a reason for postponing cost-effective measures.”<sup>87</sup> When applied to fully autonomous weapons, this principle calls for preventive action to be taken now.

---

<sup>85</sup> See Schmitt and Thurnher, “Out of the Loop,” *Harvard National Security Journal*, pp. 255-256 and 263 (arguing technology could “theoretically” be developed to assess proportionality and doubt about distinction in armed conflict situations).

<sup>86</sup> Noel Sharkey, “Computing Experts from 37 Countries Call for Ban on Killer Robots,” *International Committee for Robot Arms Control*, October 16, 2013, <http://icrac.net/2013/10/computing-experts-from-37-countries-call-for-ban-on-killer-robots/> (accessed February 16, 2014).

<sup>87</sup> Rio Declaration on Environment and Development, adopted June 14, 1992, U.N. Doc. A/CONF.151/26 (vol. 1), 31 ILM 874, 1992, principle 15. While the Rio Declaration applies the precautionary principle to environmental harms, the principle can be adapted to other contexts. See Human Rights Watch and IHRC, “The Need for New Law to Ban Fully Autonomous Weapons,” pp. 12-14.

Human Rights Watch and IHRC recommend a preemptive ban on fully autonomous weapons, which would forestall the troubling consequences described in this report and have great humanitarian benefits. It would also help prevent an arms race, block proliferation, and stop development before countries invest so heavily in this technology that they do not want to give it up.<sup>88</sup> In determining the future of fully autonomous weapons, the international community should seriously consider their human rights implications and ensure the core components of this body of law receive protection.

---

<sup>88</sup> See generally Human Rights Watch and IHRC, “The Need for New Law to Ban Fully Autonomous Weapons.”

## **Acknowledgments**

This report was researched and written by Bonnie Docherty, senior researcher in the Arms Division of Human Rights Watch and senior clinical instructor at the International Human Rights Clinic (IHRC) at Harvard Law School. Alysa Harder, Francesca Procaccini, and Caroline Sacerdote, Harvard Law students in the IHRC, contributed to the research. Steve Goose, director of the Arms Division at Human Rights Watch, edited the report, and Mary Wareham, advocacy director of the Arms Division, provided additional feedback. Dinah PoKempner, general counsel, and Tom Porteous, deputy program director, also reviewed the report.

This report was prepared for publication by Andrew Haag, associate in the Arms Division, Grace Choi, publications director, and Fitzroy Hepkins, administrative manager.

# SHAKING THE FOUNDATIONS

## The Human Rights Implications of Killer Robots

Fully autonomous weapons threaten to violate the foundational rights to life and to a remedy and to undermine the principle of human dignity that underlies human rights. These future weapons, also called “killer robots,” would identify and fire on targets without meaningful human intervention. International debate over fully autonomous weapons has focused on their role in armed conflict and their ability to comply with international humanitarian law.

*Shaking the Foundations* raises serious concerns about the use of fully autonomous weapons in law enforcement operations, in which human rights law would apply. Such weapons would be unable to evaluate the need for and proportionality of using deadly force the way human beings do. They could not be preprogrammed to handle all law enforcement scenarios. And they would lack human qualities, such as judgment and empathy, that enable police to avoid unlawful arbitrary killings in unforeseen situations.

In addition, there would be legal and practical obstacles to holding anyone—a superior officer, programmer, or manufacturer—sufficiently responsible for the conduct of a fully autonomous weapon. This accountability gap would infringe on the right to a remedy and interfere with punishment and deterrence.

Finally, fully autonomous weapons could not respect the inherent dignity and value of human beings. As inanimate machines without lives to lose, these robots would be unable truly to comprehend and weigh the significance of the deaths they cause.

The human rights implications of fully autonomous weapons compound other legal, ethical, and scientific concerns about use of these weapons. To prevent this host of problems, countries should adopt a preemptive ban on development, production, and use of fully autonomous weapons.



[hrw.org](http://hrw.org)

**IHRC**

INTERNATIONAL HUMAN RIGHTS CLINIC  
HUMAN RIGHTS PROGRAM AT HARVARD LAW SCHOOL



<http://hrp.law.harvard.edu/clinic/>