

MIND THE GAP

The Lack of Accountability for Killer Robots

HUMAN
RIGHTS
WATCH

IHRC

INTERNATIONAL HUMAN RIGHTS CLINIC
HUMAN RIGHTS PROGRAM AT HARVARD LAW SCHOOL



Mind the Gap

The Lack of Accountability for Killer Robots

Copyright © 2015 Human Rights Watch

All rights reserved.

Printed in the United States of America

ISBN: 9781623132408

Cover design by Rafael Jimenez

Cover cartoon by Russell Christian

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/>



APRIL 2015



INTERNATIONAL HUMAN RIGHTS CLINIC
HUMAN RIGHTS PROGRAM AT HARVARD LAW SCHOOL

978-1-6231-32408

Mind the Gap

The Lack of Accountability for Killer Robots

Summary	1
Recommendations	5
I. Fully Autonomous Weapons and Their Risks	6
Risks Posed by Fully Autonomous Weapons	7
International Humanitarian Law: Distinction and Proportionality	8
International Human Rights Law: Right to Life and Human Dignity	9
The Shortcomings of Regulation	10
II. The Importance of Personal Accountability	13
Purposes of Criminal Responsibility	14
International Legal Obligations	15
III. Criminal Accountability	18
Direct Responsibility	19
Indirect or Command Responsibility	20
Existence of a Crime	21
Actual or Constructive Knowledge	22
Punish or Prevent	23
Effective Control	24
Conclusion	25
IV. Civil Accountability	26
The Practical Difficulties of Civil Accountability	27
Obstacles to Military Accountability	27
Obstacles to Manufacturer Accountability	30
Immunity for Military Manufacturers	30
Loopholes in Products Liability Law	31
A No-Fault Compensation System	35
Conclusion	36
Conclusion	37
Acknowledgments	38

Summary

Fully autonomous weapons, also known as “killer robots,” raise serious moral and legal concerns because they would possess the ability to select and engage their targets without meaningful human control. Many people question whether the decision to kill a human being should be left to a machine. There are also grave doubts that fully autonomous weapons would ever be able to replicate human judgment and comply with the legal requirement to distinguish civilian from military targets. Other potential threats include the prospect of an arms race and proliferation to armed forces with little regard for the law.

These concerns are compounded by the obstacles to accountability that would exist for unlawful harm caused by fully autonomous weapons. This report analyzes in depth the hurdles to holding anyone responsible for the actions of this type of weapon. It also shows that even if a case succeeded in assigning liability, the nature of the accountability that resulted might not realize the aims of deterring future harm and providing retributive justice to victims.

Fully autonomous weapons themselves cannot substitute for responsible humans as defendants in any legal proceeding that seeks to achieve deterrence and retribution. Furthermore, a variety of legal obstacles make it likely that humans associated with the use or production of these weapons—notably operators and commanders, programmers and manufacturers—would escape liability for the suffering caused by fully autonomous weapons. Neither criminal law nor civil law guarantees adequate accountability for individuals directly or indirectly involved in the use of fully autonomous weapons.

The need for personal accountability derives from the goals of criminal law and the specific duties that international humanitarian and human rights law impose. Regarding goals, punishment of past unlawful acts aims to deter the commission of future ones by both perpetrators and observers aware of the consequences. In addition, holding a perpetrator responsible serves a retributive function. It gives victims the satisfaction that a guilty party was condemned and punished for the harm they suffered and helps avoid collective blame and promote reconciliation. Regarding duties, international humanitarian law mandates personal accountability for grave breaches, also known as war crimes. International human rights law, moreover, establishes a right to a remedy, which encompasses various

forms of redress; for example, it obliges states to investigate and prosecute gross violations of human rights law and to enforce judgments in victims' civil suits against private actors.

Existing mechanisms for legal accountability are ill suited and inadequate to address the unlawful harms fully autonomous weapons might cause. These weapons have the potential to commit criminal acts—unlawful acts that would constitute a crime if done with intent—for which no one could be held responsible.¹ A fully autonomous weapon itself could not be found accountable for criminal acts that it might commit because it would lack intentionality. In addition, such a robot would not fall within the “natural person” jurisdiction of international courts. Even if such jurisdiction were amended to encompass a machine, a judgment would not fulfill the purposes of punishment for society or the victim because the robot could neither be deterred by condemnation nor perceive or appreciate being “punished.”

Human commanders or operators could not be assigned direct responsibility for the wrongful actions of a fully autonomous weapon, except in rare circumstances when those people could be shown to have possessed the specific intention and capability to commit criminal acts through the misuse of fully autonomous weapons. In most cases, it would also be unreasonable to impose criminal punishment on the programmer or manufacturer, who might not specifically intend, or even foresee, the robot's commission of wrongful acts.²

The autonomous nature of killer robots would make them legally analogous to human soldiers in some ways, and thus it could trigger the doctrine of indirect responsibility, or command responsibility. A commander would nevertheless still escape liability in most cases. Command responsibility holds superiors accountable only if they knew or should have known of a subordinate's criminal act and failed to prevent or punish it. These criteria set a high bar for accountability for the actions of a fully autonomous weapon.

Command responsibility deals with prevention of a crime, and since robots could not have the mental state to commit an underlying crime, command responsibility would never be

¹ Any crime consists of two elements: an act and a mental state. A fully autonomous weapon could commit a criminal act (such as an act listed as an element of a war crime), but it would lack the mental state (often intent) to make these wrongful actions prosecutable crimes.

² The programmer and manufacturer might also lack the military operator's or commander's understanding of the circumstances or variables the robot would encounter and respond to, which would diminish the likelihood it could be proved they intended the unlawful act.

available in situations involving these weapons. If that issue were set aside, however, given that the weapons are designed to operate independently, a commander would not always have sufficient reason or technological knowledge to anticipate the robot would commit a specific unlawful act. Even if he or she knew of a possible unlawful act, the commander would often be unable to prevent the act, for example, if communications had broken down, the robot acted too fast to be stopped, or reprogramming was too difficult for all but specialists. In addition, “punishing” the robot after the fact would not make sense. In the end, fully autonomous weapons would not fit well into the scheme of criminal liability designed for humans, and their use would create the risk of unlawful acts and significant civilian harm for which no one could be held criminally responsible.

An alternative approach would be to hold a commander or a programmer liable for negligence if, for example, the unlawful acts brought about by robots were reasonably foreseeable, even if not intended. Such civil liability can be a useful tool for providing compensation for victims and provides a degree of deterrence and some sense of justice for those harmed. It imposes lesser penalties than criminal law, however, and thus does not achieve the same level of social condemnation associated with punishment of a crime.

Regardless of the nature of the penalties, attempts to use civil liability mechanisms to establish accountability for harm caused by fully autonomous weapons would be equally unlikely to succeed. On a practical level, even in a functional legal system, most victims would find suing a user or manufacturer difficult because their lawsuits would likely be expensive, time consuming, and dependent on the assistance of experts who could deal with the complex legal and technical issues implicated by the use of fully autonomous weapons. The legal barriers to civil accountability are even more imposing than the practical barriers. They are exemplified by the limitations of the civil liability system of the United States, a country which is generally friendly to litigation and a leader in the development of autonomous technology.

Immunity for the US military and its defense contractors presents an almost insurmountable hurdle to civil accountability for users or producers of fully autonomous weapons. The military is immune from lawsuits related to: (1) its policy determinations, which would likely include a choice of weapons, (2) the wartime combat activities of military forces, and (3) acts committed in a foreign country. Manufacturers contracted by the military are similarly immune from suit when they design a weapon in accordance with

government specifications and without deliberately misleading the military. These same manufacturers are also immune from civil claims relating to acts committed during wartime.

Even without these rules of immunity, a plaintiff would find it challenging to establish that a fully autonomous weapon was legally defective for the purposes of a product liability suit. The complexity of an autonomous robot's software would make it difficult to prove that it had a manufacturing defect, that is, a production flaw that prevented it from operating as designed. The fact that a fully autonomous weapon killed civilians would also not necessarily indicate a manufacturing defect: a robot could have acted within the bounds of international humanitarian law, or the deaths could have been attributable to a programmer who failed to foresee and plan for the situation. The plaintiffs' ability to show that the weapons' design was in some way defective would be impeded by the complexity of the technology, the unavailability of existing alternative weapons to serve as points of comparison, and the limited utility of warnings where the hazards inherent in a weapon that operates independently are generally apparent but unpredictable in specifics.

A system of providing compensation without establishing fault has been proposed for other autonomous technologies. Under such a scheme, victims would have to provide only proof that they had been harmed, not proof that the product was defective. This approach would not, however, fill the accountability gap that would exist were fully autonomous weapons used. No-fault compensation is not the same as accountability, and victims of fully autonomous weapons are entitled to a system that punishes those responsible for grave harm, deters further harm, and shows that justice has been done.

Some proponents of fully autonomous weapons argue that the use of the weapons would be acceptable in limited circumstances, but once they are developed and deployed, it would be difficult to restrict them to such situations. Proponents also note that a programmer or operator could be held accountable in certain cases, such as when criminal intent is proven. As explained in this report, however, there are many other foreseeable cases involving fully autonomous weapons where criminal and civil liability would not succeed. Even if the law adopted a strict liability regime that allowed for compensation to victims, it would not serve the purposes of deterrence and retribution that international humanitarian and human rights law seek to achieve. This report argues that states should eliminate this accountability gap by adopting an international ban on fully autonomous weapons.

Recommendations

In order to preempt the accountability gap that would arise if fully autonomous weapons were manufactured and deployed, Human Rights Watch and Harvard Law School's International Human Rights Clinic (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.

I. Fully Autonomous Weapons and Their Risks

Fully autonomous weapons are weapons systems that would select and engage targets without meaningful human control. They are also known as killer robots or lethal autonomous weapons systems. Because of their *full* autonomy, they would have no “human in the loop” to direct their use of force and thus would represent the step beyond current remote-controlled drones.³

Fully autonomous weapons do not yet exist, but technology is moving in their direction, and precursors are already in use or development. For example, many countries use weapons defense systems—such as the Israeli Iron Dome and the US Phalanx and C-RAM—that are programmed to respond automatically to threats from incoming munitions. In addition, prototypes exist for planes that could autonomously fly on intercontinental missions (UK Taranis) or take off and land on an aircraft carrier (US X-47B).⁴

The lack of meaningful human control places fully autonomous weapons in an ambiguous and troubling position. On the one hand, while traditional weapons are tools in the hands of human beings, fully autonomous weapons, once deployed, would make their own determinations about the use of lethal force. They would thus challenge long-standing notions of the role of arms in armed conflict, and for some legal analyses, they would be more akin to a human soldier than to an inanimate weapon. On the other hand, fully autonomous weapons would fall far short of being human. Indeed, they would resemble other machines in their lack of certain human characteristics, such as judgment, compassion, and intentionality. This quality underlies many of the objections that have been raised in response to the prospect of fully autonomous weapons. This report

³ In general, robotic weapons are unmanned, and they are frequently divided into three categories based on the level of autonomy and, consequentially, the amount of human involvement in their actions:

- **Human-*in*-the-Loop Weapons:** Robots that can select targets and deliver force only with a human command;
- **Human-*on*-the Loop Weapons:** Robots that can select targets and deliver force under the oversight of a human operator who can override the robots’ actions; and
- **Human-*out-of*the Loop Weapons:** Robots that are capable of selecting targets and delivering force without any human input or interaction.

In this report, the term “fully autonomous weapon” is used to refer to both “out-of-the-loop” weapons as well as weapons that allow a human on the loop, but with supervision that is so limited that the weapons are effectively “out of the loop.”

⁴ 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>, pp. 9-20.

analyzes one of the most important of these objections: the difficulty of securing accountability when fully autonomous weapons become involved in the commission of unlawful acts.

Risks Posed by Fully Autonomous Weapons

While proponents of fully autonomous weapons tout such military advantages as faster-than-human reaction times and enhanced protection of friendly forces, opponents, including Human Rights Watch and IHRC, believe the cumulative risks outweigh any benefits.⁵ From a moral perspective, many people find objectionable the idea of delegating to machines the power to make life-and-death determinations in armed conflict or law enforcement situations. In addition, although fully autonomous weapons would not be swayed by fear or anger, they would lack compassion, a key safeguard against the killing of civilians. Because these weapons would revolutionize warfare, they could also trigger an arms race; if one state obtained such weapons, other states might feel compelled to acquire them too. Once developed, fully autonomous weapons would likely proliferate to irresponsible states or non-state armed groups, giving them machines that could be programmed to indiscriminately kill their own civilians or enemy populations. Some critics also argue that the use of robots could make it easier for political leaders to resort to force because using such robots would lower the risk to their own soldiers; this dynamic would likely shift the burden of armed conflict from combatants to civilians.

Finally, fully autonomous weapons would face significant challenges in complying with international law. They would lack human characteristics generally required to adhere during armed conflict to foundational rules of international humanitarian law, such as the rules of distinction and proportionality.⁶ When used in non-armed conflict situations, such as law enforcement, fully autonomous weapons would have the potential to undermine the human right to life and the principle of human dignity. The obstacles to compliance, which are elaborated on below, not only endanger civilians, but also increase the need for an effective system of legal accountability to respond to any violations that might occur.

⁵ Human Rights Watch and IHRC, “Advancing the Debate on Killer Robots: 12 Key Arguments for a Preemptive Ban on Fully Autonomous Weapons,” May 2014, <http://www.hrw.org/news/2014/05/13/advancing-debate-killer-robots>, pp. 20-21.

⁶ According to Heather Roff, even if a robot could distinguish combatants from civilians, an additional problem that she calls the “Strategic Robot Problem” arises. Allowing a Strategic Robot to determine targeting lists and processes “undermines existing command and control structures, eliminating what little power humans have over the trajectory and consequences of war.” Heather M. Roff, “The Strategic Robot Problem: Lethal Autonomous Weapons in War,” *Journal of Military Ethics*, vol. 13, no. 3 (2014), p. 212.

International Humanitarian Law: Distinction and Proportionality

Fully autonomous weapons would face great, if not insurmountable, difficulties in reliably distinguishing between lawful and unlawful targets as required by international humanitarian law. The weapons would lack human qualities that facilitate making such determinations, particularly on contemporary battlefields where combatants often seek to conceal their identities. Distinguishing an active combatant from a civilian or injured or surrendering soldier requires more than the deep sensory and processing capabilities that might be developed. It also depends on the qualitative ability to gauge human intention, which involves interpreting subtle, context-dependent clues, such as tone of voice, facial expressions, or body language. Humans possess the unique capacity to identify with other human beings and are thus equipped to understand the nuances of unforeseen behavior in ways in which machines—which must be programmed in advance—simply are not.

The obstacles presented by the principle of distinction are compounded when it comes to proportionality, which prohibits attacks in which expected civilian harm outweighs anticipated military advantage. Because proportionality relies heavily on a multitude of contextual factors, the lawful response to a situation could change considerably by slightly altering the facts. According to the US Air Force, “proportionality in attack is an inherently subjective determination that will be resolved on a case-by-case basis.”⁷ It would be nearly impossible to pre-program a machine to handle the infinite number of scenarios it might face. In addition, international humanitarian law depends on human judgment to make subjective decisions, and proportionality is ultimately “a question of common sense and good faith for military commanders.”⁸ It would be difficult to replicate in machines the judgment that a “reasonable military commander” exercises to assess proportionality in unforeseen or changing circumstances.⁹ Non-compliance with the principle of proportionality, in addition to the failure to distinguish between civilians and combatants, could lead to an unlawful loss of life.

⁷ US Air Force Judge Advocate General’s Department, “Air Force Operations and the Law: A Guide for Air and Space Forces,” first edition, 2002, p. 27.

⁸ International Committee of the Red Cross (ICRC), *Commentary on the Additional Protocols of 8 June 1977 to the Geneva Conventions of 12 August 1949*, (Geneva: Martinus Nijhoff Publishers, 1987), pp. 679, 682.

⁹ The generally accepted standard for assessing proportionality is whether a “reasonable military commander” would have launched a particular attack. International Criminal Tribunal for the Former Yugoslavia (ICTY), “Final Report to the Prosecutor by the Committee Established to Review the NATO Bombing Campaign Against the Federal Republic of Yugoslavia,” June 8, 2000, <http://www.icty.org/sid/10052> (accessed May 8, 2014), para. 50.

International Human Rights Law: Right to Life and Human Dignity

Fully autonomous weapons have the potential to contravene the right to life, which is the bedrock of international human rights law. According to the International Covenant on Civil and Political Rights (ICCPR), “No one shall be arbitrarily deprived of his life.”¹⁰ The use of lethal force 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 specific circumstance. In addition, when encountering unforeseen situations, fully autonomous weapons would be prone to carrying out arbitrary killings because they would face challenges in meeting the three requirements for the use of force. 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.¹¹

The concept of human dignity also lies at the heart of international human rights law. The opening words of the Universal Declaration of Human Rights (UDHR) assert, “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.”¹² In ascribing inherent dignity to all human beings, the UDHR, drawing on Kant, implies that everyone has worth that deserves respect.¹³ Fully autonomous weapons would possess the power to kill people yet, because they are not human, they would be unable to respect their victims’ dignity. As inanimate machines, they could comprehend neither the value of individual human life nor the significance of its loss. Therefore, on top of

¹⁰ 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).

¹¹ For more information, see Human Rights Watch and IHRC, *Shaking the Foundations: The Human Rights Implications of Killer Robots*, May 2014, <http://hrw.org/node/125251>.

¹² Universal Declaration of Human Rights (UDHR), adopted December 10, 1948, G.A. Res. 217A(III), U.N. Doc. A/810, pmb., para. 1 (1948) (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.”

¹³ Jack Donnelly, “Human Dignity and Human Rights,” in Swiss Initiative to Commemorate the 60th Anniversary of the UDHR, *Protecting Dignity: Agenda for Human Rights*, June 2009, p. 10. See also *ibid.*, p. 21 (quoting Immanuel Kant’s *The Metaphysics of Morals*, which states, “Man regarded as a person ... is exalted above any price; ... he is not to be valued merely as a means ... he possesses a dignity (absolute inner worth) by which he exacts respect for himself from all other rational beings in the world.”).

putting civilians at risk, allowing fully autonomous weapons to make determinations to take life away would conflict with the principle of dignity.¹⁴

The Shortcomings of Regulation

Some proponents of fully autonomous weapons argue that the answer to the legal concerns discussed above is to limit the circumstances in which the weapons are used. They contend that there are some potential uses, no matter how limited or unlikely, where fully autonomous weapons would be both militarily valuable and capable of conforming to the requirements of international humanitarian law. One proponent, for example, notes that “[n]ot every battlespace contains civilians.”¹⁵ Other proponents maintain that fully autonomous weapons could be used lawfully under “limited circumstances,” such as in attacks on “nuclear-tipped mobile missile launchers, where millions of lives were at stake.”¹⁶ These authors generally favor restricting the use of fully autonomous weapons to specific types of locations or purposes.¹⁷ Regulations could come in the form of a legally binding instrument or a set of gradually developed, informal standards.¹⁸

The regulatory approach does not eliminate all the risks of fully autonomous weapons. It is difficult to restrict use of weapons to narrowly constructed scenarios. Once fully autonomous weapons came into being under a regulatory regime, they would be vulnerable to misuse. Even if regulations restricted use of fully autonomous weapons to certain locations or specific purposes, after the weapons entered national arsenals

¹⁴ 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, A/HRC/23/47, April 9, 2013, http://www.ohchr.org/Documents/HRBodies/HRCouncil/RegularSession/Session23/A-HRC-23-47_en.pdf (accessed February 14, 2014), p. 20.

¹⁵ Michael N. Schmitt, “Autonomous Weapon Systems and International Humanitarian Law: A Reply to the Critics,” *Harvard National Security Journal Features* (2013), <http://harvardnsj.org/wp-content/uploads/2013/02/Schmitt-Autonomous-Weapon-Systems-and-IHL-Final.pdf> (accessed April 1, 2015), p. 11.

¹⁶ Paul Scharre, “Reflections on the Chatham House Autonomy Conference,” post to “Lawfare” (blog), March 3, 2014, <http://www.lawfareblog.com/2014/03/guest-post-reflections-on-the-chatham-house-autonomy-conference/> (accessed April 20, 2014).

¹⁷ See, for example, Armin Krishnan, “Automating War: The Need for Regulation,” *Contemporary Security Policy*, vol. 30, no. 1 (2009), p. 188.

¹⁸ Kenneth Anderson and Matthew Waxman, “Law and Ethics for Autonomous Weapon Systems: Why a Ban Won’t Work and How the Laws of War Can,” *Jean Perkins Task Force on National Security and Law* (2013), http://media.hoover.org/sites/default/files/documents/Anderson-Waxman_LawAndEthics_r2_FINAL.pdf (accessed March 31, 2015), p. 22 (explaining that “[b]y ‘international norms’ here, we do not mean new binding legal rules only—whether treaty rules or customary international law—but instead the gradual fostering of widely-held expectations about legally or ethically appropriate conduct, whether formally binding or not”).

countries that usually respect international humanitarian law could be tempted in the heat of battle or in dire circumstances to use the weapons in ways that increased the risk of laws of war violations. For example, before adoption of the 2008 Convention on Cluster Munitions, proponents of cluster munitions often maintained that the weapons could be lawfully launched on a military target alone in an otherwise unpopulated desert. Even generally responsible militaries, however, made widespread use of cluster munitions in populated areas. Such theoretical possibilities should not be used to legitimize weapons, including fully autonomous ones, that pose significant humanitarian risks when used in less exceptional situations.

The existence of fully autonomous weapons would also make possible their acquisition by repressive regimes or non-state armed groups that might disregard the restrictions or alter or override any programming designed to regulate a robot's behavior. They could use the weapons in intentional or indiscriminate attacks against their own people or civilians in other countries with horrific consequences.

An absolute, legally binding ban on fully autonomous weapons would provide several distinct advantages over formal or informal constraints. It would maximize protection for civilians in conflict because it would be more comprehensive than regulation. It would be more effective as it would prohibit the existence of the weapons and be easier to enforce.¹⁹ A ban could have a powerful stigmatizing effect, creating a widely recognized new standard and influencing even those that did not join a treaty. Finally, it would obviate other problems with fully autonomous weapon, such as moral objections and the potential for an arms race.

A ban would also minimize the problems of accountability that come with regulation. By legalizing limited use of fully autonomous weapons, regulation would open the door to situations where accountability challenges arise. If the weapons were developed and deployed, there would be a need to hold persons responsible for violations of international law involving the use of these weapons. Even if responsibility were assigned

¹⁹ Enforcement of regulations on fully autonomous weapons, as on all regulated weapons, could be challenging and leave room for error, increasing the potential for harm to civilians. Instead of knowing that any use of fully autonomous weapons is unlawful, countries, international organizations, and nongovernmental organizations would have to monitor the use of the weapons and determine in every case whether use complied with the regulations. There would probably be debates about enforcement and the scope of the regulations—for example, what constituted a populated area, where use of certain weapons might be restricted.

via a strict liability scheme, it would merely produce compensation and neither reflect moral judgment nor achieve accountability's goals of deterrence and retribution. The rest of this report elaborates on the hurdles to ensuring accountability for unlawful acts committed by fully autonomous weapons that meets these goals.

II. The Importance of Personal Accountability

When violations of international law occur, perpetrators should not be allowed to escape accountability. As noted international law scholar M. Cherif Bassiouni states, “Impunity for international crimes and for systematic and widespread violations of fundamental human rights is a betrayal of our human solidarity with the victims of conflicts to whom we owe a duty of justice, remembrance, and compensation.”²⁰ In the same vein, the Basic Principles and Guidelines on the Right to a Remedy and Reparation, adopted by the United Nations in 2005, recognize the importance of “the international legal principles of accountability, justice and the rule of law” and lay out the elements of an accountability regime.²¹ Accountability can come in the form of state responsibility, which may lead to changes in a country’s conduct.²² This report, however, focuses on personal accountability, for either natural or legal persons, which punishes the conduct of a specific offender rather than the state.

The purpose of assigning personal responsibility under criminal law is to deter future violations and to provide retribution to victims. Such accountability is required by international law: international humanitarian law imposes a duty to prosecute war crimes, and international human rights law establishes a right to a remedy for infringement of other rights. Human rights law also promotes civil liability, which can in practice meet some of the goals of criminal responsibility but carries less moral weight. While state responsibility for the unlawful acts of fully autonomous weapons could be assigned relatively easily to the user state, as discussed in Chapters 3 and 4, it would be difficult to ascribe personal responsibility for those acts. Understanding the underlying goals and legal obligations associated with accountability is a prerequisite to grasping the significance of the potential gap in responsibility for humans or corporations.

²⁰ M. Cherif Bassiouni, “Accountability for Violations of International Humanitarian Law and Other Serious Violations,” <http://www.sos-attentats.org/publications/bassiouni.violations.pdf> (accessed April 1, 2015), p. 422.

²¹ UN General Assembly, “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), Resolution 60/47, March 21, 2006, A/Res/60/147, pmbl.

²² See International Law Commission, “Draft Articles on the Responsibility of States for Internationally Wrongful Acts,” 2001, http://legal.un.org/ilc/texts/instruments/english/draft%20articles/9_6_2001.pdf (accessed March 23, 2015).

Purposes of Criminal Responsibility

One of the primary reasons to hold individuals accountable is to deter harmful behavior. If individuals are punished for unlawful acts, they may be less likely to repeat them. Holding offenders responsible can also discourage future infractions by other actors, who fear being punished in the same way. According to Dinah Shelton, author of the treatise *Remedies in International Human Rights Law*, “[d]eterrence ... is assumed to work because rational actors weigh the anticipated costs of transgressions against the anticipated benefits.”²³ She adds that “[d]eterrence literature also shows a correlation between the certainty of consequences and the reduction of offences.”²⁴ For deterrence to have the maximum effect, potential offenders must have advance notice of the prospect of accountability so that they can consider the consequences before they act. Public assurances that steps are being taken to diminish the likelihood of new offenses can also provide consolation to victims and society.

Accountability serves an additional retributive function.²⁵ The commission of an unlawful act against another person “conveys a message that the victim’s rights are not sufficiently important to refrain from violating them in pursuit of another goal,” while punishment shows “criminals and others that they wronged the victim and thus implicitly recognizes the victim’s plight and honors the victim’s moral claims.”²⁶ Holding an individual accountable gives victims the satisfaction of knowing that someone is being condemned and punished for the harm they suffered, and it sends the message that the lives and rights of victims have value. By specifying who is most proximately responsible, personal accountability also avoids collective blame, which can spur revenge or impede reconciliation, and plays a vital role in post-conflict resolution, both for the victims and for the community as a whole.²⁷

²³ Dinah Shelton, *Remedies in International Human Rights Law* (Oxford: Oxford University Press, 2005), p. 12.

²⁴ *Ibid.*

²⁵ *Ibid.*

²⁶ *Ibid.*, pp. 12-13.

²⁷ Assigning responsibility to specific offenders provides the benefits of discouraging collective punishment and promoting reconciliation. Following conflict, there is often an inclination to punish one side as a group. Accountability for individual perpetrators can, however, “promote peace and reconciliation and reduce the prospect of future violations by breaking the collective cycle of guilt.” Allison Marston Danner and Jenny S. Martinez, “Guilty Associations: Joint Criminal Enterprise, Command Responsibility, and the Development of International Criminal Law,” *California Law Review*, vol. 93 (2005), p. 93. Antonio Cassese, former judge at the International Criminal Tribunal for the Former Yugoslavia, explained, “[T]rials establish individual responsibility over collective assignation of guilt, i.e., they establish that not all Germans were responsible for the Holocaust.” In addition, he noted that “victims are prepared to be reconciled with their erstwhile tormentors because they know that the latter have now paid for their crimes.” Antonio Cassese, “Reflections on International Criminal Justice,” *Modern Law Review*, vol. 61 (1998), p. 6.

Although not strictly speaking a goal of accountability, a third purpose of criminal law is compensatory justice. Providing compensation to victims through a legal process aims to rectify the harm experienced and thereby restore the victim to the condition he or she was in before the harm was inflicted.²⁸ Such compensation seeks to restore not only an economic balance between the parties, but also a moral balance. In addition, it furthers the autonomy of victims by providing them with funds that may assist them to achieve their individual life goals that were stymied by the initial injury.²⁹ While compensation has several valuable functions, however, it is not a substitute for deterrence and retribution.

International Legal Obligations

International humanitarian law and international human rights law both require accountability for legal violations. International humanitarian law establishes a duty to prosecute criminal acts committed during armed conflict. The Fourth Geneva Convention and its Additional Protocol I, the key legal instruments on civilian protection, oblige states to prosecute “grave breaches,” i.e., war crimes, such as willfully targeting civilians or launching an attack with the knowledge it would be disproportionate.³⁰ The obligation links international humanitarian law with international criminal law; the former “provides the source of many of the crimes [prosecuted under] ... international criminal law.”³¹ As a result, many international criminal tribunals, including the International Criminal Court (ICC), the International Criminal Tribunal for the former Yugoslavia (ICTY), and the International Criminal Tribunal for Rwanda (ICTR), have jurisdiction over war crimes as well as genocide and crimes against humanity.³² While prosecutions can also take place at the domestic level, the personal accountability mandated by international humanitarian law is often implemented through international criminal law.

²⁸ Shelton, *Remedies in International Human Rights Law*, p. 10.

²⁹ *Ibid.*, p. 11.

³⁰ 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 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. 85-86.

³¹ Danner and Martinez, “Guilty Associations: Joint Criminal Enterprise, Command Responsibility, and the Development of International Criminal Law,” *California Law Review*, p. 80.

³² Rome Statute of the International Criminal Court (Rome Statute), A/CONF.183/9, July 17, 1998, entered into force July 1, 2002, art. 1; Statute of the International Tribunal for the Prosecution of Persons Responsible for Violations of International Humanitarian Law Committed in the Territory of the Former Yugoslavia Since 1991 (ICTY Statute), adopted by Security Council May 25, 1993, U.N. Doc. S/Res/827, art. 1; Statute of the International Criminal Tribunal for Rwanda (ICTR Statute), adopted by Security Council November 8, 1994, U.N. Doc. S/Res/955, art. 1.

International human rights law establishes the right to a remedy for abuses of all human rights. 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.”³³ Highlighting the value of this right for deterrence, the Human Rights Committee has stated that, “the purposes of the Covenant [ICCPR] would be defeated without an obligation ... to take measures to prevent a recurrence of a violation.”³⁴ Under the right to a remedy, international human rights law, like international humanitarian law, mandates prosecution of individuals for serious violations of the law, notably genocide and crimes against humanity. According to the Human Rights Committee, 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.³⁵ A failure to investigate and, where appropriate, prosecute “could in and of itself give rise to a separate breach of the Covenant.”³⁶

The 2005 Basic Principles and Guidelines on the Right to a Remedy and Reparation articulate in one document the obligations on states to provide effective avenues for accountability under both international humanitarian law and international human rights law. Adopted by the UN General Assembly, these principles lay out a victim’s right to a remedy, which encompasses a state’s duty to investigate and prosecute. The Basic Principles and Guidelines specifically require states to punish individuals who are found guilty of serious violations of either body of international law.³⁷ The inclusion of this obligation demonstrates the importance the international community places on personal accountability.

³³ ICCPR, art. 2(3). See also 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.”); African [Banjul] Charter on Human and People’s 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. 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, adopted November 22, 1969, O.A.S. Treaty Series No. 36, 1144 U.N.T.S. 123, entered into force July 18, 1978, 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, 213 U.N.T.S. 222, entered into force September 3, 1953, art. 13 (“Everyone whose rights and freedoms as set forth in this Convention are violated shall have an effective remedy before a national authority.”).

³⁴ UN Human Rights Committee, General Comment No. 31, Nature of the General Legal Obligations of States Parties to the Covenant, U.N. Doc. CCPR/C/21/Rev.1/Add.13 (2004), para. 17.

³⁵ *Ibid.*, paras. 15 and 18.

³⁶ *Ibid.*

³⁷ 2005 Basic Principles and Guidelines, art. 4.

The right to a remedy and accountability in general, however, are not limited to criminal prosecution.³⁸ The Basic Principles and Guidelines “are without prejudice to the right to a remedy and reparation” for all violations of international human rights and humanitarian law, not just those serious enough to rise to the level of crimes.³⁹ The standards encourage redress through civil law, notably by requiring states to enforce judgments related to private claims brought by victims against individuals or entities.⁴⁰ Options for accountability thus extend beyond criminal prosecution to domestic civil litigation. The obstacles to both options as a response to the unlawful actions of fully autonomous weapons will be discussed in the following chapters.

³⁸ For example, states can provide remedies in the form of reparations including “restitution, compensation, rehabilitation, satisfaction and guarantees of non-repetition.” 2005 Basic Principles and Guidelines, arts. 18-23.

³⁹ Ibid., art. 26

⁴⁰ Ibid., art. 17.

III. Criminal Accountability

Criminal accountability is a key tool for punishing wrongful acts of the past and deterring those of the future. The international community has established such accountability for the gravest crimes—war crimes, crimes against humanity, and genocide—most notably through *ad hoc* tribunals responding to the atrocities in the former Yugoslavia and Rwanda and the permanent International Criminal Court (ICC). Explaining the importance of criminal accountability, the Nuremberg Tribunal, the post-World War II predecessor to these courts, stated: “Crimes against international law are committed by *men*, not by abstract entities, and only by punishing individuals who commit such crimes can the provisions of international law be enforced.”⁴¹

Robots are not men, however, and the deployment of fully autonomous weapons could represent a step backward for international criminal law. The use of such weapons would create the potential for a vacuum of personal legal responsibility for the type of civilian harm associated with war crimes or crimes against humanity. Gaps in criminal accountability for fully autonomous weapons would exist under theories of both direct responsibility and indirect responsibility (also known as command responsibility).

Any crime consists of two elements. There must be a criminal act, the *actus reus*, and the act must be perpetrated with a certain mental state, or *mens rea*. Fully autonomous weapons could commit criminal acts, a term this report uses for actions that would fall under the *actus reus* element. For example, a fully autonomous weapon would have the potential to direct attacks against civilians, kill or wound a surrendering combatant, or launch a disproportionate attack, all of which are elements of war crimes under the Rome Statute of the ICC.⁴² By contrast, fully autonomous weapons could not have the mental state required to make these wrongful actions crimes; because they would not have moral agency, they would lack the independent intentionality that must accompany the commission of criminal acts to establish criminal liability. Since robots could not satisfy both elements of a crime, they could not be held legally responsible, and one must look to the operator or commander as alternative options for accountability. Unless an operator or

⁴¹ “Judgment of the Nuremberg International Military Tribunal,” *American Journal of International Law*, vol. 41 (1947), http://www.loc.gov/rr/frd/Military_Law/pdf/NT_Vol-I.pdf (accessed March 22, 2015), p. 223 (emphasis added).

⁴² Rome Statute, art. 8.

commander acted with criminal intent or at least knowledge of the robot’s criminal act, however, there would be significant obstacles to holding anyone responsible for a fully autonomous weapon’s conduct under international criminal law.

Direct Responsibility

Direct responsibility holds offenders liable for playing an active role in the commission of a crime. Direct responsibility creates accountability for the direct perpetrator—the one who pulls the trigger; it also covers other actors who are directly involved because, for example, they planned or ordered a crime.⁴³ Robots could not themselves be held responsible for their actions under this doctrine for three reasons. First, as noted above, although they might commit a criminal act, they could not have the mental state required to perpetrate a crime. Second—and closely related to this first point—international criminal tribunals generally limit their jurisdiction to “natural persons,” that is, human beings, because they have intentionality to commit crimes.⁴⁴ Third, even if this jurisdiction were expanded, on a practical level, fully autonomous weapons could not be punished because they would be machines that could not experience suffering or apprehend or learn from punishment.⁴⁵ Thus, fully autonomous weapons would present a novel accountability gap: the entity selecting and engaging targets—which to date has always been a human—could not be held directly responsible for a criminal action that resulted from the unlawful selection or engagement of targets. With no direct responsibility applicable to fully autonomous weapons, there would be no accountability for the actual perpetrator of the criminal acts causing civilian harm.

Furthermore, there would be insufficient direct responsibility for a human who deployed or operated a fully autonomous weapon that committed a criminal act.⁴⁶ A gap could arise because fully autonomous weapons by definition would have the capacity to act autonomously and therefore could launch independently and unforeseeably an indiscriminate attack against civilians or those *hors de combat*. In such a situation,

⁴³ For example, the ICTY Statute allows direct responsibility for a “person who planned, instigated, ordered, committed, or otherwise aided and abetted in the planning, preparation and execution of a crime.” ICTY Statute, art. 7.

⁴⁴ Rome Statute, art. 25 (“The Court shall have jurisdiction over natural persons pursuant to this Statute.”); ICTY Statute, art. 6.

⁴⁵ Robert Sparrow, “Killer Robots,” *Journal of Applied Philosophy*, vol. 24, no. 1 (2007), p. 72. See also Heather M. Roff, “Killing in War: Responsibility, Liability and Lethal Autonomous Robots,” in *Routledge Handbook of Ethics and War*, eds. Fritz Allhof, Nicholas G. Evans, and Adam Henschke (New York: Routledge, 2013), p. 355 (arguing that fully autonomous weapons “lack this capacity for moral autonomy, and thus responsibility.”).

⁴⁶ While this section focuses on the challenges of establishing responsibility of an operator or commander, the analysis is the same regarding the responsibility of a programmer or manufacturer.

regardless of whether the fully autonomous weapon were conceived of as a subordinate soldier under a human commander or as a weapon being employed by a human operator, direct responsibility would likely not attach to a human commander or operator. Using a commander-subordinate analogy, the commander would not be *directly* responsible for the robot's specific actions since he or she did not order them.⁴⁷ At best, the commander or operator would only be responsible for deploying the robot, and liability would rest on whether that decision under the circumstances amounted to an intention to commit an indiscriminate attack. Some have argued that the more apt analogy is that of an operator using a non-autonomous weapon, but that position does not change the analysis. It would still be difficult to ascribe direct responsibility because, in this scenario, the operator is unlikely to have foreseen that the weapon would cause civilian casualties. Under international criminal law, a human could be *directly* responsible for criminal acts committed by a robot only if he or she deployed the robot intending to commit a crime, such as willfully killing civilians.⁴⁸

Even if direct responsibility were legally possible, there would be evidentiary challenges to proving accountability. Robots would have at least two parties providing the equivalent of orders: the operator and the programmer (and there would often be many individuals involved in programming). Each party could try to shift blame to the others in an attempt to avoid responsibility. Therefore, proving which party was responsible for the orders that led to the targeting of civilians might prove difficult, even if a user did intentionally employ fully autonomous weapons to commit a crime.

Indirect or Command Responsibility

In international criminal law, indirect responsibility, also known as command responsibility, holds a military commander or civilian superior criminally liable for failing to prevent or punish a subordinate's crime.⁴⁹ Specifically, liability under command

⁴⁷ The challenges of holding the commander *indirectly* responsible are examined below in the discussion of command responsibility.

⁴⁸ This direct liability would attach regardless of how one viewed the human-robot relationship. Under the commander-subordinate analogy, just as a soldier's selection of particular targets does not absolve a commander who ordered that soldier to kill civilians, a robot's independent selection of particular targets would not absolve the human who deployed it from criminal liability. Under the operator-traditional weapon analogy, the operator would be criminally responsible if he or she intended to use the fully autonomous weapon against civilians.

⁴⁹ Historically, "command responsibility" was limited to military commanders, but since World War II the doctrine has also been applied to civilian superiors, leading some sources to use the more inclusive term "superior responsibility." "Command responsibility" and "superior responsibility" are sometimes used interchangeably today. This document uses the traditional

responsibility occurs when a superior fails to take necessary and reasonable measures to prevent or punish the criminal acts of a subordinate over whom the superior has effective control, once that superior knows or has reason to know of the criminal acts.⁵⁰ Command responsibility holds the superior accountable for dereliction of duty, a crime of omission.⁵¹ Prosecuting command responsibility is difficult because it typically requires state cooperation and provision of internal military evidence to prove the elements of knowledge and effective control.⁵²

Significant obstacles would exist to establishing accountability for criminal acts committed by fully autonomous weapons under the doctrine of command responsibility. Because of their ability to make independent determinations about selecting and engaging targets, fully autonomous weapons would be analogous to subordinate soldiers in standard command responsibility analysis. Thus, a commander would theoretically be liable under command responsibility if he or she knew or had reason to know that the robot would commit or had committed a crime, failed to prevent or punish the robot, and had effective control over the robot. Achieving this type of criminal accountability would be legally challenging for at least four reasons.

Existence of a Crime

First, command responsibility only arises when a subordinate commits a chargeable criminal offense.⁵³ The subordinate must satisfy all elements of the underlying crime, not merely attempt the crime or commit some other inchoate offense such as conspiracy.⁵⁴ As discussed above, robots cannot satisfy the *mens rea* element of a crime, and thus cannot be charged with a crime even if they commit criminal acts. As a result, fully autonomous

term “command responsibility” because fully autonomous weapons will likely be developed and primarily deployed by professional militaries.

⁵⁰ *Prosecutor v. Delalić*, ICTY, Case No. IT-96-21-T, Judgment (Trial Chamber), Nov. 16, 1998, para. 346. Although the above description accurately captures current customary international law on command responsibility, the doctrine has been codified slightly differently in various international agreements. Protocol I, arts. 86-87; ICTY Statute, art. 7(3); ICTR Statute, art. 6(3); Statute of the Special Court for Sierra Leone (SCSL Statute), adopted by Security Council August 14, 2000, U.N. Doc. S/Res/13145, art. 6(3); Rome Statute, art. 28.

⁵¹ *Prosecutor v. Halilović*, ICTY, Case No. IT-01-48-T, Judgment (Trial Chamber), Nov. 16, 2005, para. 54.

⁵² Christine Bishai, “Superior Responsibility, Inferior Sentencing: Sentencing Practice at the International Criminal Tribunals,” *Northwestern Journal of International Human Rights*, vol. 11 (2013), pp. 83, 86-87.

⁵³ Guénaél Mettraux, *The Law of Command Responsibility* (Oxford: Oxford University Press, 2009) pp. 131-134.

⁵⁴ *Ibid.*

weapons as “subordinates” cannot commit underlying crimes for which commanders could be held accountable under command responsibility.

Actual or Constructive Knowledge

Second, even if a criminal act committed by a robot were considered sufficient for the command responsibility doctrine in the case of fully autonomous weapons, the doctrine would be ill suited for these weapons because it would be difficult for commanders to acquire the appropriate level of knowledge. Command responsibility is only triggered if a commander has actual or constructive knowledge of the crime, that is, the commander must know or have a reason to know of the crime. Actual knowledge of an impending criminal act would only occur if a fully autonomous weapon communicated its target selection prior to initiating an attack. If a human operator exercised oversight, effective review, and veto power over a robot’s specific targets and strikes, however, then such a robot would have a “human on the loop” and would not constitute the type of *fully* autonomous weapon being discussed here. Assuming a fully autonomous weapon would be subject to less than complete oversight of its targeting decisions, it is uncertain exactly how frequently commanders would have actual knowledge of an impending criminal strike, much less a real opportunity to override the strike.

Constructive knowledge would likely be more relevant to the situation of fully autonomous weapons. The command responsibility standard for constructive knowledge requires commanders to have information that puts them “on notice of the risk” of a subordinate’s crime that is “sufficiently alarming to justify further inquiry.”⁵⁵ Such sufficiently alarming information triggers a duty to investigate, and failure to do so can lead to liability under command responsibility. Commanders cannot be held liable for negligently failing to find out information without having received some alarming information.⁵⁶ Actual knowledge of past offenses by a particular set of subordinates may constitute sufficiently alarming information to necessitate further inquiry, and thus may constitute constructive knowledge (reason to know) of future criminal acts satisfying the *mens rea* of command

⁵⁵ *Prosecutor v. Strugar*, ICTY, Case No. IT-01-42-A, Judgment (Appeals Chamber), July 17, 2008, paras. 297-98.

⁵⁶ *Prosecutor v. Blaškić*, ICTY, Case No. IT-95-14-A, Judgment (Appeals Chamber), July 29, 2004, paras. 62, 406.

responsibility.⁵⁷ The analysis is heavily fact-specific, however, and considers the specific circumstances of the commander at the time in question.⁵⁸

This standard raises a number of questions about what constitutes constructive knowledge in the context of fully autonomous weapons. For example, would knowledge of past unlawful acts committed by one robot provide notice of risk only for that particular robot, or for all robots of its make, model, and/or programming? Would knowledge of one type of past unlawful act, such as a robot's mistaking of a civilian for a combatant (a failure of distinction), trigger notice of the risk of other types of unlawful acts, such as a failure to accurately determine the proportionality of a future strike? Would fully autonomous weapons be predictable enough to provide commanders with the requisite notice of potential risk? Would liability depend on a particular commander's individual understanding of the complexities of programming and autonomy? Depending on the answers to these questions, a commander might escape liability for the acts of a fully autonomous weapon. Once a commander is considered on notice, he or she would have to take "necessary and reasonable" measures to prevent foreseeable criminal activity by those robots.⁵⁹ There would be no command responsibility for the robots' criminal acts until that point, however, because it would be unjust to hold commanders liable for criminal acts that they could not prevent or punish due to genuine lack of knowledge. The uncertainties of the knowledge standard in the context of fully autonomous weapons would make it difficult to apply command responsibility.

Punish or Prevent

A third obstacle to the application of command responsibility is the requirement that a commander can punish or prevent a crime. Robots cannot be punished, making one of the omissions criminalized under the principle of command responsibility irrelevant in the context of criminal acts involving the use of fully autonomous weapons. Command responsibility could, therefore, only arise with respect to failure to *prevent* criminal acts by fully autonomous weapons, but finding commanders accountable under the obligation to prevent would have challenges. As just discussed, a commander might not have adequate

⁵⁷ *Prosecutor v. Strugar*, ICTY, paras. 299-301.

⁵⁸ *Ibid.*, para. 299.

⁵⁹ ICTY Statute, art. 7(3); Protocol 1, art. 87(3) (requiring commanders "to initiate such steps as are *necessary to prevent* such violations of the Conventions or this Protocol, and, where appropriate, to initiate disciplinary or penal action against violators thereof") (emphasis added).

knowledge of the criminal act to trigger the duty to prevent it. The very nature of autonomy would make it difficult to predict a robot's next attack in changing circumstances. In addition, commanders might lack the actual ability to prevent fully autonomous weapons from committing criminal acts. For example, a key advantage of autonomy—the ability to make calculations and initiate attack or response more quickly than human judgment can—would make interrupting a criminal act particularly difficult, even assuming a robot possessed the safeguard of having its ability to attack dependent on maintaining a communications link with a human operator.⁶⁰ Finally, judges with little knowledge of autonomy, robotics, or complex programming might be hesitant to question a commander's claim that an autonomous weapon's future behavior was too unforeseeable to merit criminal liability and/or that the commander took all reasonable measures to prevent criminal acts from occurring.

Effective Control

Fourth, command responsibility would require commanders to have effective control over fully autonomous weapons. The “material ability to control the actions of subordinates is the touchstone of individual [command] responsibility,”⁶¹ specifically the “material ability to prevent or punish criminal conduct.”⁶² As discussed above, prevention in the case of fully autonomous weapons would be difficult, and punishment impossible. The fully autonomous weapon's fast processing speed as well as unexpected circumstances, such as communication interruptions, programming errors, or mechanical malfunctions, might prevent commanders from being able to call off an attack. Furthermore, a commander's formal responsibility for a subordinate—or in this case, for a fully autonomous weapon—would not guarantee command responsibility for that weapon's criminal acts without *de facto* control.⁶³ For any criminal acts committed during a period without effective control, there would be no command responsibility.

⁶⁰ Furthermore, the fully autonomous weapon might be designed to continue to operate autonomously if it were already deployed and there were a breakdown in communications. In fact, continuing to operate autonomously should communications break down is often cited as an advantage of these types of weapons. In such a situation, however, the robot might commit a criminal act that cannot be prevented. The commander could, of course, prevent a criminal act by choosing not to deploy the fully autonomous weapon, but doing so when no crime was foreseeable would undermine the advantages of having the technology.

⁶¹ *Prosecutor v. Kayishema*, ICTR, Case No. ICTR-95-1-T, Judgment (Trial Chamber II), May 21, 1999, para. 229.

⁶² *Prosecutor v. Delalić et al.*, Case No. IT-96-21-A, Judgment (Appeals Chamber), Feb. 20, 2001, para. 256 (known as the *Čelebići* case).

⁶³ *Ibid.*, para. 197 (“[T]he possession of *de jure* power in itself may not suffice for the finding of command responsibility if it does not manifest in effective control.”).

Conclusion

The doctrine of command responsibility could fall short for many reasons, including insufficient knowledge, an inability to prevent or punish, and a lack of effective control. Given the absence of direct responsibility for a fully autonomous weapon's criminal acts, when command responsibility fails, criminal accountability would not be available at all. The accountability gap would be especially problematic before there is notice of a fully autonomous weapon's criminal acts. When a human soldier commits a crime, there is accountability even if commanders do not have constructive knowledge because the actual perpetrator can be held directly responsible. If a fully autonomous weapon commits a criminal act, by contrast, neither the robot nor its commander could be held accountable for crimes that occurred before the commander was put on notice. During this accountability-free period, a robot would be able to commit repeated criminal acts before any human had the duty or even the ability to stop it.

The challenges in ensuring criminal responsibility undermine the aims of accountability. If there were no consequences for human operators or commanders, future criminal acts could not be deterred, and victims of fully autonomous weapons would likely view themselves as targets of preventable attacks for which no one was condemned and punished. The inadequate deterrence and retribution under existing international criminal law raises serious questions about the wisdom of producing and using fully autonomous weapons.

IV. Civil Accountability

Civil accountability often serves as an alternative or supplement to criminal accountability. Under civil law, which focuses on harms against individuals rather than society, victims instead of prosecutors bring suits. Monetary damages are the most common penalty, and compensation, along with stigmatization of the guilty party, can help victims feel a sense of justice and deter future wrongful acts. The consequences of breaching obligations under civil law, however, are arguably less severe than under criminal law because they do not include imprisonment. Civil damages lack the social condemnation associated with criminal accountability.

In addition to concerns over the adequacy of civil accountability, there would be significant practical and legal obstacles to holding either the user or manufacturer of a fully autonomous weapon liable under this body of law.⁶⁴ If a party designed or manufactured a weapon with the specific intent to kill civilians, it would likely be held legally responsible, at least under criminal law. But situations that do not involve a clear unlawful intent would be more challenging. While civil law often deals with cases where the defendant was reckless or negligent, it would not be feasible for most victims of fully autonomous weapons to bring a civil suit against a user or manufacturer. Furthermore, the military and its contractors are largely immune under civil law, at least in some countries, and product liability suits are unlikely to succeed against them.⁶⁵

The following analysis deals primarily with the US civil regime, but that regime is especially relevant to the case of fully autonomous weapons. The United States, which has been described as a “land of opportunity” for litigation, is often perceived as possessing the most plaintiff-friendly tort regime in the world.⁶⁶ If victims could not effectively make use of

⁶⁴ The term “manufacturers” is used here to refer to the individuals or organizations involved in programming the software enabling fully autonomous weapons to act autonomously, as well as those organizations involved in manufacturing the weapon in its final form. These parties could all be potential targets of lawsuits by victims of fully autonomous weapons or their families.

⁶⁵ Fully autonomous weapons themselves could not be held liable under civil law, regardless if they were seen as analogous to a human soldier. Even if possible, such accountability would be meaningless because weapons could not pay damages if found guilty.

⁶⁶ Charles F. Rysavy and Pranita A. Raghavan, “The (Often Insurmountable) Hurdles Facing Foreign Claimants Prosecuting Class Actions in American Courts,” *Tort Trial & Insurance Practice Law Journal*, vol. 42 (2006), p. 1; See also Stephanie A. Scharf and Traci M. Braun, “Foreign Plaintiffs Battle to Keep Class Claims in U.S. Courts,” *FDL’s Update* (January-February 2007), p. 32.

US civil accountability mechanisms, it is unlikely that they would be more successful in other jurisdictions. Furthermore, the United States and US manufacturers are among the leaders in the development of the autonomous technology that would lead to fully autonomous weapons. Any effective and comprehensive civil accountability regime would therefore need to apply in a US context.

The Practical Difficulties of Civil Accountability

The victims of fully autonomous weapons or their relatives could face hurdles in suing either the users or manufacturers of these weapons, even in a functional legal system. Regardless of whether legal action is undertaken in the United States or a different jurisdiction, lawsuits can be lengthy and expensive and require legal and technical experts. The barriers of time, money, and expertise are often sufficient to deter litigation in a purely domestic context. Those barriers would be even greater for a victim living far from the country that used the fully autonomous weapon, or from the headquarters of the company that manufactured it. A lawyer or nongovernmental organization could assist victims in litigating some cases and find experts to analyze the technology. The government, however, would likely respond that the technology was a state secret or impose other obstacles to interfere with an analysis of defects.

Obstacles to Military Accountability

A civilian victim of an unlawful act committed by a fully autonomous weapon could potentially sue the military force that used the weapon. For example, the relative of someone killed by such a weapon could seek redress for a wrongful death on the basis that the military negligently or recklessly used a fully autonomous weapon that was prone to violations of international humanitarian law or used it in specific situations where it was likely to cause civilian casualties.⁶⁷ In civil litigation, victims are far more apt to sue the military authority than the individual soldier operating the weapon. The military possesses a greater financial capacity than the individual to provide compensation, and at least in the US context, government employees are immune from civil suit except where they are accused of violating the US Constitution or the suit is authorized by statute.⁶⁸

⁶⁷ For a full discussion of the inability of fully autonomous weapons to comply with international humanitarian law, see Human Rights Watch and IHRC, “Advancing the Debate on Killer Robots,” pp. 4-8.

⁶⁸ Federal Employees Liability Reform and Tort Compensation Act of 1988, Pub. L. 100-694, title 28, November 18, 1988, 28 U.S.C. § 2679.

Even with this shift toward state responsibility, however, victims would probably not prevail in their pursuit of accountability. The US military would likely be legally immune from all such suits relating to its decisions regarding fully autonomous weapons, and the doctrine of sovereign immunity similarly protects other governments in relation to decisions on weaponry ordered or used, especially in foreign combat situations.⁶⁹

The US government is presumptively immune from civil suits.⁷⁰ It has waived this immunity in some circumstances, most notably in the Federal Torts Claims Act (FTCA). The waiver is subject to a few exceptions, however, three of which are particularly relevant when dealing with the military: (1) the discretionary function exception; (2) the combatant activities exception; and (3) the foreign country exception.⁷¹ If a case falls under any of these exceptions, then the government is immune from suit.

The discretionary function exception means, in essence, that government agencies cannot be sued for actions taken while implementing the government's policy goals.⁷² Acts such as the selection of military equipment designs or the choice to use a fully autonomous weapon in a particular environment could be covered by this exception.⁷³ US courts are often reluctant to inquire too deeply into the government's foreign affairs policies, including the mechanics of military decision making.⁷⁴ They would therefore probably apply this exception if they were faced with a case initiated by the victim of a fully autonomous weapon.

The combatant activities exception immunizes the government from civil claims relating to the wartime combat activities of military forces—the most likely context for the use of fully

⁶⁹ See, for example, *Just v. British Columbia*, [1989] 2 SCR 1228 (finding the Canadian government immune from suit regarding its policy decisions); Crown Proceedings Act 1947, Section 10 (stating that the UK government and soldiers themselves are immune for all actions taken by members of the armed forces on duty); *Jurisdictional Immunities of the State (Germany v. Italy)*, International Court of Justice, Judgment, February 3, 2012 (a finding by the International Court of Justice that states are immune even for civil suits relating to serious violations of international law).

⁷⁰ John Copeland Nagle, "Waiving Sovereign Immunity in an Age of Clear Statement Rules," *Wisconsin Law Review*, vol. 1995, (1995), pp. 776-777.

⁷¹ 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; 28 U.S.C. §§ 2860(a), (j), (k) (2014).

⁷² Finkelman, "Suing the Hired Guns," *Brooklyn Journal of International Law*, p. 405.

⁷³ See, for example, *Boyle v. United Technologies Corp.*, 487 U.S. 500, 511 (1988).

⁷⁴ "Developments in the Law, Access to the Courts, The Political Question Doctrine, Executive Deference, and Foreign Relations," *Harvard Law Review*, vol. 122 (2009), p. 1193.

autonomous weapons.⁷⁵ Courts have defined the exception to include armed conflict outside a formally declared war and any activities connected to hostilities, not just armed activities themselves.⁷⁶ For example, the heirs of the deceased passengers of Iran Air Flight 655, an Iranian civilian aircraft that was shot down by the warship *USS Vincennes*, were precluded from suing the United States due to sovereign immunity under the combatant activities exception.⁷⁷

The foreign country exception bars individuals from raising claims against the US government “arising in a foreign country.”⁷⁸ This exception has been interpreted broadly to restrict civil claims relating to conduct occurring overseas, even if the activities were planned by members of the US government in the United States.⁷⁹ For example, this provision prevented the estate of an engineer killed in a take-off crash on a US military airbase in Canada from suing the US government for compensation.⁸⁰ This exception would sharply limit claims based on the deployment of fully autonomous weapons by US forces in overseas engagements.

Thus the US military would likely enjoy immunity for any claims relating to the development or use of fully autonomous weapons because they would involve policy decisions and/or wartime conduct or because they took place overseas. Even jurisdictions with highly restrictive immunity doctrines, such as the United Kingdom, might well preclude these types of suits. The British Supreme Court recently and controversially allowed the families of British soldiers killed in Iraq to sue the British government for negligence and human rights violations. The court, however, stated that the UK Ministry of Defense would still be immune for “high-level policy decisions ... or decisions made in the heat of battle.”⁸¹

⁷⁵ 28 U.S.C. § 2680(j).

⁷⁶ Finkelman, “Suing the Hired Guns,” *Brooklyn Journal of International Law*, p. 425; *Koohi v. United States*, 976 F.2d 1328, 1334 (9th Cir. 1992); *Johnson v. United States*, 170 F.2d 767, 770 (9th Cir.1948).

⁷⁷ *Koohi*, 976 F.2d at 1337.

⁷⁸ 28 U.S.C. § 2680(k).

⁷⁹ *Sosa v. Alvarez-Machain*, 542 U.S. 692, 710 (2004); David W. Fuller, “Intentional Torts and Other Exceptions to the Federal Tort Claims Act,” *University of St. Thomas Law Journal*, vol. 8 (2011), p. 383.

⁸⁰ *United States v. Spelar*, 338 U.S. 217, 219 (1949).

⁸¹ “Iraq Damages Cases: Supreme Court Rules Families Can Sue,” *BBC*, June 19, 2013, <http://www.bbc.com/news/uk-22967853> (accessed March 21, 2015).

Obstacles to Manufacturer Accountability

If a fully autonomous weapon killed or injured a non-combatant as a result of production, programming, or design flaws, one method of achieving at least partial accountability would be to hold the manufacturer that created and programmed the weapon liable for its actions. Doing so, however, would be extremely difficult under existing US law. Military contractors often enjoy immunity from civil suit. In addition, products liability law, the most common method of imposing civil liability on manufacturers, cannot adequately accommodate claims regarding autonomous devices, which are capable of making independent determinations.

Immunity for Military Manufacturers

Military contractors, like the military itself, are usually immune from litigation in US courts as an extension of the government's own immunity for two reasons. First, US courts have found military contractors immune in cases involving a weapon whose design the government chose. Manufacturers cannot be held liable for any harm caused by a defective weapon if: (1) the government approved particular and precise specifications for the weapon, (2) the weapon conformed to those specifications, and (3) the manufacturer did not deliberately fail to inform the government of a known danger regarding the weapon of which the government was unaware.⁸² This legal rule barred, for example, a suit against the manufacturer of a helicopter that crashed, killing a serviceman.⁸³

As long as the military was actively involved in the development of fully autonomous weapons, and did not merely “rubber stamp” contractor design decisions or buy weapons with a standardized design off the shelf, manufacturers would likely escape liability for their role in providing the weapons to the military.⁸⁴ Because fully autonomous weapons would be highly complex and adapted to specific situations, militaries would almost certainly play an important part in their design and development, and their manufacturers would consequently be immune from being sued.⁸⁵

⁸² *Boyle*, 487 U.S. at 512.

⁸³ *Ibid.*

⁸⁴ Colin P. Cahoon, “Boyle Under Siege,” *Journal of Air Law and Commerce*, vol. 59 (1994), p. 835. This position has been adopted in several US circuits. See, for example, *Trevino v. Gen. Dynamics Corp.*, 865 F.2d 1474, 1480 (5th Cir. 1989) (“We hold that ‘approval’ under the *Boyle* defense requires more than a rubber stamp.”). If the manufacturer of a fully autonomous weapon built a weapon that did not conform to military specifications or if it deliberately failed to inform the government of known risks, however, this immunity would not apply and civil accountability might be possible.

⁸⁵ The US Department of Defense, for example, has issued a number of policy documents describing their plans regarding the development of autonomous weapons. See US Department of Defense, “Unmanned Integrated Systems Roadmap,” 2013, <http://www.defense.gov/pubs/DOD-USRM-2013.pdf> (accessed March 21, 2015); U.S. Department of Defense, “Directive

Second, US military contractors, like the military itself, are immune from claims arising from wartime activities. For example, the manufacturer of the air-defense weapon that shot down the Iranian civilian aircraft in the *Vincennes* incident mentioned above was subject to immunity because the case involved a “combatant activity.”⁸⁶ Similarly, the manufacturer of a missile that accidentally targeted friendly forces on the Kuwait-Saudi border was immunized under this exception.⁸⁷ For the same reason, manufacturers would likely escape liability for the use of fully autonomous weapons in combat.

Loopholes in Products Liability Law

In theory, products liability law would offer another avenue to civil accountability for the manufacturers of fully autonomous weapons.⁸⁸ This area of civil law holds entities engaged in the business of designing, producing, selling, or distributing products liable for harm to persons or property caused by defects in those products.⁸⁹

Products liability law was not developed to cope with autonomous technology. Even if military contractors were not immune under the rules discussed above, plaintiffs would have great difficulty succeeding in a product liability suit against the manufacturers of fully autonomous weapons. Victims of a fully autonomous weapon could conceivably sue the weapon’s manufacturer arguing that a manufacturing or design defect existed, but neither defect theory could be easily relied on in fully autonomous weapons cases.⁹⁰

Manufacturing Defects

A manufacturing defect refers to a product’s failure to meet its design specifications.⁹¹ For example, if a car’s steering wheel is faulty and seizes up or a microwave spontaneously combusts, a manufacturing defect was likely present because the products presumably did not act as designed. Manufacturers of fully autonomous weapons would be subject to

3000.09: Autonomy in Weapons Systems,” November 21, 2012, <http://www.dtic.mil/whs/directives/corres/pdf/300009p.pdf> (accessed March 21, 2015).

⁸⁶ *Koohi*, 976 F.2d at 1330-1331.

⁸⁷ *Bentzlin v. Hughes Aircraft Co.*, 833 F. Supp. 1486 (C.D. Cal. 1993).

⁸⁸ In this case, fully autonomous weapons would be treated as analogous to a machine not a human agent such as a soldier.

⁸⁹ *Restatement (Third) of Torts: Products Liability* (1998), sec. 1.

⁹⁰ Manufacturing defect cases impose strict liability, meaning that no fault is required. Design defect cases impose a more complicated type of liability, sometimes called “strict products liability,” which “achieve[s] the same general objectives as does liability predicated on negligence.” In neither case is a plaintiff required to prove the harm was intended. *Ibid.*, secs. 1-2.

⁹¹ *Ibid.*, sec. 2(a).

this type of claim if a plaintiff argued that the weapon failed to operate in the way intended by the manufacturer. There are two major difficulties in applying the doctrine of manufacturing defects to autonomous systems.

First, this doctrine is generally used in the case of tangible defects (like the microwave explosion) and appears not to have been applied to alleged defective software because “nothing tangible is manufactured.”⁹² Yet it is precisely the software-based, independent decision-making capacity of fully autonomous weapons that would be most concerning from both a legal and ethical perspective.

Second, even if plaintiffs were not necessarily barred from making claims regarding software-based manufacturing defects, it would be difficult for them to prove the existence of such a defect in a fully autonomous weapon. These types of cases typically involve plaintiffs asking the court to infer a manufacturing defect based on a product’s failure, without proof of the specific defect itself.⁹³ For example, if a car’s brakes fail, a court may use that fact to infer the existence of a manufacturing defect even if the plaintiff cannot point to a particular problem with the brakes because one can normally assume that brakes functioning as designed would not fail.⁹⁴

In the case of fully autonomous weapons, however, the mere fact that such a weapon killed civilians would be insufficient to prove that the weapon failed and was not acting according to design.⁹⁵ Even if a fully autonomous weapon were designed to provide appropriate military responses that complied with international humanitarian law, that body of law allows the incidental killing of known civilians in certain situations, such as when an attack is discriminate (that is, it targets military objectives) and proportionate (that is, the military advantage outweighs likely civilian harm). The killing of a civilian might thus be consistent with adequate performance in both manufacture and design. Alternatively, the killing could be due either to a malfunction in programming or hardware or to a flawed design that failed to provide a legally compliant response in a given

⁹² Jeffrey K. Gurney, “Sue My Car Not Me: Products Liability and Accidents Involving Autonomous Vehicles,” *University of Illinois Journal of Law, Technology and Policy*, vol. 2013 (2013), p. 259.

⁹³ Andrew P. Garza, “‘Look Ma, No Hands!’: Wrinkles and Wrecks in the Age of Autonomous Vehicles,” *New England Law Review*, vol. 46 (2012), p. 591.

⁹⁴ Kevin Funkhouser, “Paving the Road Ahead: Autonomous Vehicles, Products Liability, and the Need for A New Approach,” *Utah Law Review* (2013), pp. 447-448.

⁹⁵ For an application of a similar criticism with respect to autonomous vehicles, see *ibid.*

situation. Armed conflict can present very fluid, unpredictable, and complex circumstances, and in developing fully autonomous weapons, programmers and manufacturers might not predict every situation such a weapon might face. It would therefore be difficult to discern malfunction alone from a failure to program or design for the myriad contingencies that require legal compliance.

Because of these problems—the inherent difficulty in pointing to specific defects in non-physical software and the difficulty of inferring a defect from a malfunction—manufacturing defect claims would not be viable options in the case of fully autonomous weapons.

Design Defects

A design defect involves a product design that presents a foreseeable and unreasonable risk of harm that could have been reduced or avoided with a reasonable alternative design.⁹⁶ In this type of case, the plaintiff would claim that the very design of fully autonomous weapons was inherently defective, perhaps because of the inability of these weapons adequately to comply with international humanitarian law. Like manufacturing defect claims, these types of lawsuits would be unlikely to succeed. There are two major approaches to design defects: one focused upon the expectations of consumers and the other upon the risks associated with a product. Neither would be well suited to claims arising from the actions of autonomous machines.

Under the consumer expectations test, plaintiffs can demonstrate the existence of a design defect by asserting that a particular design is inconsistent with ordinary consumers' expectations of safety.⁹⁷ This type of analysis is sometimes used for technologies with which ordinary consumers are extremely familiar, but it is generally disfavored when dealing with technology that is sufficiently complicated that consumers do not, according to courts, have "reasonable expectations" regarding its capabilities.⁹⁸ For example, US courts have found technology such as airbags or cruise control to be too complicated for this standard to apply.⁹⁹ Considering that even the most basic autonomous weapon would

⁹⁶ *Restatement (Third) of Torts: Products Liability*, sec. 2.

⁹⁷ Douglas A. Kysar, "The Expectations of Consumers," *Columbia Law Review*, vol. 103 (2003), pp. 1701-1702.

⁹⁸ Garza, "'Look Ma, No Hands!'," *New England Law Review*, p. 600; Terrence F. Kiely and Bruce L. Ottley, *Understanding Products Liability Law* (San Francisco: LexisNexis, 2006), pp. 138-139.

⁹⁹ Garza, "'Look Ma, No Hands!'," *New England Law Review*, pp. 598-599.

be far more complex and sophisticated than an automobile cruise control system, victims would presumably face significant obstacles in pursuing this approach.

Under the risk-utility test, the plaintiff demonstrates the existence of a design defect by showing the risks posed by a product outweigh the benefits.¹⁰⁰ When analyzing whether a manufacturer should be liable for a design defect, US courts can take into account the possible harms a product could cause, the existence of any safer alternative designs in the same category, and/or the adequacy of the warning of a product's risks to the consumer.¹⁰¹ Applying these factors to fully autonomous weapons is problematic. The manufacturer might assert that the technology benefits rather than harms the public and the military by reducing the threat to friendly forces. This balance, however, is not the one required by international humanitarian law, which emphasizes the protection of civilians from unnecessary harm rather than force protection. At least until the technology becomes widely available there would likely be few similar products with which to compare the weapons in question. In addition, the risks to human life involved in using deadly weapons are so obvious it could be difficult to argue that general warnings were necessary, and more specific warnings might not be useful given the unpredictability inherent in fully autonomous weapons. The complexity of autonomous systems would also create a practical bar to this type of case because plaintiffs would need to hire expensive expert witnesses to testify about alternative designs.¹⁰²

In addition to facing difficulties in applying the two tests, courts might have policy reasons to refrain from holding the manufacturers of fully autonomous weapons liable under the design defect theory. When dealing with ordinary products, courts have found manufacturers rather than mass-market consumers responsible because manufacturers are seen as being in the best position to ensure that a product is not defective.¹⁰³ After all, they are the ones who design and produce the product. This conclusion would be far less reasonable or intuitive when dealing with fully autonomous weapons made by a military contractor for a highly sophisticated military buyer that specifies anticipated uses.¹⁰⁴ These

¹⁰⁰ Kysar, "The Expectations of Consumers," *Columbia Law Review*, p. 1702.

¹⁰¹ Garza, "'Look Ma, No Hands!'," *New England Law Review*, p. 603; Funkhouser, "Paving the Road Ahead," *Utah Law Review*, p. 457.

¹⁰² Gurney, "Sue My Car Not Me," *University of Illinois Journal of Law, Technology and Policy*, pp. 247, 263-264.

¹⁰³ Guido Calabresi and Jon T. Hirschoff, "Toward a Test for Strict Liability in Torts," *Yale Law Journal*, vol. 81 (1972), p. 1060.

¹⁰⁴ See Richard J. Heafey and Don M. Kennedy, *Product Liability: Winning Strategies and Techniques* (New York: Law Journal Press, 2015), sec. 4.10.

weapons would presumably be custom-produced for government clients through specific procurement negotiations and contract requirements. In this case, courts would likely view the appropriate target for a lawsuit as the military authorities who commissioned and set out the specifications for the weapon, rather than the manufacturer who was largely following instructions. For similar reasons, US courts have not allowed suits against manufacturers for custom-manufactured trains¹⁰⁵ or custom-designed factory equipment.¹⁰⁶

A No-Fault Compensation System

Given the difficulties in applying civil law to complex autonomous systems, one alternative would be to adopt a no-fault compensation scheme for fully autonomous weapons. Unlike a products liability regime, it would require only proof of harm, not proof of defect. Victims could thus receive financial or in-kind assistance similar to that awarded in other civil suits without having to overcome the evidentiary hurdles related to proving a defect. Such no-fault systems are often used where a sometimes highly dangerous product or activity is nevertheless deemed socially valuable; they facilitate employment of the risky but useful product by providing compensation to victims, establishing some predictability, and setting limits on the defendant's costs. This type of no-fault system has been used to compensate people injured by vaccines.¹⁰⁷ It has also been proposed in the context of autonomous self-driving cars, which share many of the same legal barriers to suit as fully autonomous weapons.¹⁰⁸

Under such a no-fault scheme, victims or their families would file a claim with the government that used a fully autonomous weapon that caused civilian harm and receive compensation. They would not need to show negligence, recklessness, or fault for the harm; so long as the harm occurred, compensation would be due.

There are several problems with a no-fault scheme for fully autonomous weapons. First, it is difficult to imagine many governments that would be willing to put such a legal regime

¹⁰⁵ *Beckles v. General Electric Corp.*, 670 N.Y.S.2d 539 (N.Y. App. Div. 1998).

¹⁰⁶ *Estep v. Rieter Automotive North America, Inc.* 774 N.E.2d 323 (Ohio App. 2002).

¹⁰⁷ Funkhouser, "Paving the Road Ahead," *Utah Law Review*, pp. 460-461 (describing the US National Childhood Vaccination Injury Act).

¹⁰⁸ *Ibid.*, pp. 458-459; Julie Goodrich, "Driving Miss Daisy: An Autonomous Chauffeur System," *Houston Law Review*, vol. 51 (2013), p. 284.

into place.¹⁰⁹ Second, it is by no means clear that the social value of fully autonomous weapons would outweigh their apparent risks and justify a regime of this kind. These machines would be designed to make independent decisions on targets and even kill some civilians in the process, and when they malfunctioned, they might target civilians or cause disproportionate or unnecessary civilian harm. Their “social value” would rest on issues of cost or force protection, which is not a goal of international humanitarian law, or on the unprovable and unlikely prospect that robots would outperform human beings in the extremely difficult moral and analytic determinations required by international humanitarian law. Third, while such a no-fault scheme might assist those harmed by fully autonomous weapons, and would therefore help meet an important compensatory goal, it would do little to provide for meaningful accountability. Compensating a victim for harm is different from assigning accountability, which entails deterrence, moral blame, and the recognition by society and the offender of a victim as someone who has been wronged. By its very nature, a no-fault scheme is more focused on policy goals that are important, but limited substitutes for the meaningful system of accountability demanded by both international law and moral principle.¹¹⁰

Conclusion

Civil accountability mechanisms could not fill the accountability gap caused by the failure of traditional criminal law standards to cope with the advent of fully autonomous weapons. Due to the extensive immunity granted to the military and its contractors as well as the challenges posed by products liability law, relying on civil suits to fill the accountability gap associated with fully autonomous weapons would be impractical, unrealistic, and legally uncertain. Furthermore, even if someone were found civilly liable, the accountability that resulted would not substitute for criminal accountability in terms of deterrence, retributive justice, or moral stigma.

¹⁰⁹ Some countries have established ex gratia schemes to provide compensation to civilian victims of military operations. These schemes differ from the ones discussed here because the military gives assistance to the victims on a voluntary basis and makes explicit that it is not accepting legal fault. They provide interesting precedent for a fully autonomous weapons compensation scheme because they seek to help civilian victims of armed conflict, but because they are more discretionary than other no-fault schemes, they do even less to accomplish deterrence and retribution. See, for example, United States Government Accountability Office, “Military Operations: The Department of Defense’s Use of Solatia and Condolence Payments in Iraq and Afghanistan,” <http://www.gao.gov/assets/270/261104.pdf> (accessed March 26, 2015); Matthias Gebauer, “Aftermath of an Afghanistan Tragedy: Germany to Pay \$500,000 for Civilian Bombing Victims,” *Spiegel Online*, August 6, 2010, <http://www.spiegel.de/international/germany/aftermath-of-an-afghanistan-tragedy-germany-to-pay-500-000-for-civilian-bombing-victims-a-710439.html> (accessed March 26, 2015).

¹¹⁰ For example, militaries may give no-fault compensation in part to win the hearts-and-minds of the population in the battle zone.

Conclusion

The hurdles to accountability for the production and use of fully autonomous weapons under current law are monumental. The weapons themselves could not be held accountable for their conduct because they could not act with criminal intent, would fall outside the jurisdiction of international tribunals, and could not be punished. Criminal liability would likely apply only in situations where humans specifically intended to use the robots to violate the law. In the United States at least, civil liability would be virtually impossible due to the immunity granted by law to the military and its contractors and the evidentiary obstacles to products liability suits.

While proponents of fully autonomous weapons might imagine entirely new legal regimes that could provide compensation to victims, these regimes would not capture the elements of accountability under modern international humanitarian and human rights law. For example, a no-fault regime might provide compensation, but since it would not assign fault, it would not achieve adequate deterrence and retribution or place moral blame. Because these robots would be designed to kill, someone should be held legally and morally accountable for unlawful killings and other harms the weapons cause. The obstacles to assigning responsibility under the existing legal framework and the no-fault character of the proposed compensation scheme, however, would prevent this goal from being met.

The gaps in accountability are particularly disturbing as they can create “[a] climate of impunity ... [and] can leave serious negative consequences on individual survivors and ultimately on society as a whole.”¹¹¹ The limitations on assigning responsibility thus add to the moral, legal, and technological case against fully autonomous weapons and bolster the call for a ban on their development production, and use.

¹¹¹ Shelton, *Remedies in International Human Rights Law*, p. 11.

Acknowledgments

This report was written and edited by Bonnie Docherty, senior researcher in the Arms Division of Human Rights Watch and senior clinical instructor at Harvard Law School's International Human Rights Clinic (IHRC). Avery Halfon and Dean Rosenberg, students in IHRC, contributed significantly to the research, analysis, and writing of the report. Additional assistance was provided by IHRC students Torry Castellano, Yuanmei Lu, Amit Parekh, and Danae Paterson. Steve Goose, director of the Arms Division at Human Rights Watch, edited the report. Dinah PoKempner, general counsel, and Tom Porteous, deputy program director, also reviewed the report.

External reviewers Heather Roff and Dinah Shelton provided input on the report, although the report does not necessarily reflect their views.

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

MIND THE GAP

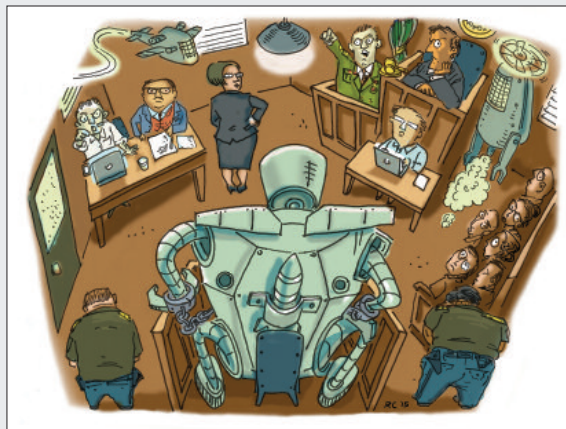
The Lack of Accountability for Killer Robots

Mind the Gap details the significant hurdles to assigning personal accountability for the actions of fully autonomous weapons, or “killer robots.” This accountability gap stems from the fact the weapons would select and engage targets without meaningful human control, and is one of several important reasons why a ban is urgently needed.

Military commanders or operators could be found guilty under criminal law if they deployed a fully autonomous weapon with the intent to commit a crime. But they would likely elude justice in the more common situation in which they could not foresee or prevent an autonomous robot’s unlawful act—such as targeting civilians, even if no human commander or programmer intended for the robot to do so.

The obstacles to accountability would be equally high under civil law. Civil liability would be virtually impossible, at least in the United States, due to the immunity granted by law to the military and its contractors and the evidentiary obstacles to products liability suits. Many other countries have similar systems of sovereign immunity.

Even if successful, a civil suit would have limited effectiveness as a tool for accountability. While monetary damages can assist victims, they are not a substitute for criminal accountability in terms of deterrence, retributive justice, and moral stigma.



The next generation of weapons in military arsenals could be “killer robots,” machines that would select and destroy specific targets without further human intervention. But if a robot broke the law, who would be held responsible? Would the programmer, manufacturer, military commander, or robot end up in court?

©2015 Russell Christian for Human Rights Watch

HUMAN
RIGHTS
WATCH

IHRC

INTERNATIONAL HUMAN RIGHTS CLINIC
HUMAN RIGHTS PROGRAM AT HARVARD LAW SCHOOL

hrw.org

www.law.harvard.edu/programs/hrp/ihr.html