

2013

At War with the Robots: Autonomous Weapon Systems and the Martens Clause

Tyler D. Evans

Follow this and additional works at: <http://scholarlycommons.law.hofstra.edu/hlr>



Part of the [Law Commons](#)

Recommended Citation

Evans, Tyler D. (2013) "At War with the Robots: Autonomous Weapon Systems and the Martens Clause," *Hofstra Law Review*: Vol. 41: Iss. 3, Article 8.

Available at: <http://scholarlycommons.law.hofstra.edu/hlr/vol41/iss3/8>

This document is brought to you for free and open access by Scholarly Commons at Hofstra Law. It has been accepted for inclusion in Hofstra Law Review by an authorized administrator of Scholarly Commons at Hofstra Law. For more information, please contact lawcls@hofstra.edu.

NOTE

AT WAR WITH THE ROBOTS: AUTONOMOUS WEAPON SYSTEMS AND THE MARTENS CLAUSE

I. INTRODUCTION

'[A] robot may not injure a human being, or, through inaction, allow a human being to come to harm.'

– Isaac Asimov, *The First Law of Robotics*¹

The Three Laws of Robotics (“Three Laws”)² are an elegant set of hierarchical rules that ethically and physically govern Isaac Asimov’s science fiction robots.³ The Three Laws are programmatically embedded in the robots’ “positronic brains,”⁴ and control their behavior and reasoning primarily to safeguard the human beings they were built to serve.⁵ The Three Laws—successful in fiction for their simplicity,

1. ISAAC ASIMOV, *Runaround*, in I, ROBOT 25, 37 (Bantam Dell 2008) [hereinafter I, ROBOT].

2. *Id.* The Three Laws are organized thusly:

“The First Law . . . ‘A robot may not injure a human being or, through inaction, allow a human being to come to harm.’ The Second Law: ‘A robot must obey the orders given it by human beings except where such orders would conflict with the First Law.’ The Third Law: ‘A robot must protect its own existence, as long as such protection does not conflict with the First or Second Law.’”

ISAAC ASIMOV, *FOUNDATION AND EARTH* 484-85 (Bantam Dell 2004) [hereinafter *FOUNDATION AND EARTH*]. The character Gregory Powell explained the hierarchical function of the Three Laws in the short story *Runaround*, stating:

“The conflict between the various rules is ironed out by the different positronic potentials in the brain. We’ll say that a robot is walking into danger and knows it. The automatic potential that Rule 3 sets up turns him back. But suppose you *order* him to walk into that danger. In that case, Rule 2 sets up a counterpotential higher than the previous one and the robot follows orders at the risk of existence.

. . . .

According to Rule 1, a robot can’t see a human come to harm because of his own inaction. Two and 3 can’t stand against it.”

I, ROBOT, *supra* note 1, at 37, 42.

3. See I, ROBOT, *supra* note 1, at 37.

4. *Id.* at ix-x, 37; see also *FOUNDATION AND EARTH*, *supra* note 2, at 485.

5. Cf. I, ROBOT, *supra* note 1, at 42. Asimov’s robots were typically slaves to their human creators. See Aaron Saenz, *The Myth of the Three Laws of Robotics - Why We Can’t Control*

novelty, and literary purposes—are ill suited for the contemporary military reality, and are generally regarded as an inadequate basis for machine ethics.⁶

Autonomous weapon systems (“AWS”), still in their technological infancy,⁷ have recently come under heavy fire from non-government organizations (“NGOs,” singularly “NGO”) and various experts who oppose the research, development, and eventual deployment of such weapon systems in armed conflicts.⁸ Specifically, Human Rights Watch (“HRW”),⁹ in conjunction with Harvard Law School’s International Human Rights Clinic (“IHRC”),¹⁰ published *Losing Humanity: The Case Against Killer Robots* (“*Losing Humanity*,” or “Report”) on November 19, 2012.¹¹ The Report contends that autonomous weapons—which do not yet exist—will be incompatible with principles of the Law of Armed Conflict (“LOAC”)¹² including distinction, proportionality, military necessity, and the Martens Clause (“Martens Clause,” or “Clause”); therefore, the Report concludes, AWS should be preemptively banned.¹³

On November 21, 2012, only three days after the publication of *Losing Humanity*, the U.S. Department of Defense (“DoD”) released

Intelligence, SINGULARITY HUB (May 10, 2011, 8:10 AM), <http://www.singularityhub.com/2011/05/10/the-myth-of-the-three-laws-of-robotics-why-we-cant-control-intelligence/> (describing Asimov’s fictional robots as “a slave race”) [hereinafter Saenz, *Myth of Three Laws*].

6. RONALD ARKIN, GOVERNING LETHAL BEHAVIOR IN AUTONOMOUS ROBOTS 48 (2009); see also Saenz, *Myth of Three Laws*, *supra* note 5.

7. Benjamin Wittes, *Does Human Rights Watch Prefer Disproportionate and Indiscriminate Humans to Discriminating and Proportionate Robots?*, LAWFARE (Dec. 1, 2012, 10:19 AM), <http://www.lawfareblog.com/2012/12/does-human-rights-watch-prefer-disproportionate-and-indiscriminate-humans-to-discriminating-and-proportionate-robots>.

8. See HUMAN RIGHTS WATCH & INT’L HUMAN RIGHTS CLINIC, HARVARD LAW SCH., *LOSING HUMANITY: THE CASE AGAINST KILLER ROBOTS* 1-2, 5 (2012), available at http://www.hrw.org/sites/default/files/reports/arms1112ForUpload_0_0.pdf [hereinafter *LOSING HUMANITY*].

9. Michael N. Schmitt, *Autonomous Weapon Systems and International Humanitarian Law: A Reply to the Critics*, HARV. NAT’L SEC. J. 1 (Feb. 5, 2013), <http://harvardnsj.org/wp-content/uploads/2013/02/Schmitt-Autonomous-Weapon-Systems-and-IHL-Final.pdf> [hereinafter Schmitt, *Autonomous Weapon Systems*]. HRW is a sophisticated NGO which operates in the field of international humanitarian law and human rights. *Id.*

10. See *International Human Rights Clinic*, HUM. RTS. PROGRAM, HARV. L. SCH., <http://www.law.harvard.edu/programs/hrp/ihr.html> (last visited July 18, 2013). IHRC, part of Harvard Law School’s Human Rights Program, promotes human rights and related law. *Id.*

11. *LOSING HUMANITY*, *supra* note 8; *Arms: New Campaign to Stop Killer Robots*, HUMAN RIGHTS WATCH (Apr. 23, 2013), <http://www.hrw.org/news/2013/04/23/arms-new-campaign-stop-killer-robots>.

12. INT’L COMM. OF THE RED CROSS, *What Is International Humanitarian Law?* (2004), available at http://www.icrc.org/eng/assets/files/other/what_is_ihl.pdf. The LOAC, interchangeably known as the Law of War and International Humanitarian Law, is part of the larger body of international law and governs the actions of belligerents in armed conflicts. *Id.*

13. *LOSING HUMANITY*, *supra* note 8, at 30-36.

Directive Number 3000.09 on Autonomy in Weapon Systems (“Directive”), which establishes in detail DoD policy for the development and use of autonomous and semi-autonomous weapon systems.¹⁴ The Directive characterizes the technical distinction between autonomous and semi-autonomous weapon systems, declares when human supervision of such systems will be required, and outlines DoD procedure for the development, review, acquisition, and use of autonomous and semi-autonomous weapon systems alike.¹⁵

AWS will ultimately be investigated and developed by states because advances in technology have historically been a major factor in determining the outcomes of wars.¹⁶ Such weapon systems, once developed, may offer more than an offensive military advantage: AWS may also be more compliant with the LOAC than human combatants.¹⁷ Stated differently, AWS may be a superior offensive means of delivering force, and a more humane means of doing so.¹⁸

Before addressing the premature claims of AWS’ opponents, the state of the technology should be made explicit: truly autonomous weapons do not yet exist.¹⁹ Critics of these weapons argue that future robotics technologies, which cannot be predicted with accuracy, will fail to comply with LOAC principles.²⁰ These allegations are, at best, based “on unfounded assumptions as to the nature of [future autonomous weapon] systems.”²¹ That there is no factual basis to examine non-existent technologies, the determination that such technologies will not

14. U.S. DEP’T OF DEFENSE, DIRECTIVE NO. 3000.09: AUTONOMY IN WEAPONS SYSTEMS 1-2 (2012), available at <http://www.dtic.mil/whs/directives/corres/pdf/300009p.pdf> [hereinafter DIRECTIVE NO. 3000.09].

15. See *id.* at 1-3.

16. See David Bell, *In Defense of Drones: A Historical Argument*, NEW REPUBLIC (Jan. 27, 2012), <http://www.newrepublic.com/article/politics/100113/obama-military-foreign-policy-technology-drones>; see also ARKIN, *supra* note 6, at 29 (“The trend is clear: Warfare will continue and autonomous robots will ultimately be deployed in its conduct.”); Michael N. Schmitt & Jeffrey S. Thurnher, “*Out of the Loop*”: *Autonomous Weapon Systems and the Law of Armed Conflict*, 4 HARV. NAT’L SEC. J. 231, 279 (2013) (“Autonomous technology is poised to revolutionize warfare. Nations will undoubtedly clamor for the weapon systems that this technology will make possible.”).

17. ARKIN, *supra* note 6, at 29-30.

18. See *id.* at 29; see also Wittes, *supra* note 7 (“To call for a *per se* ban on autonomous weapons is to insist as a matter of [international law] on preserving a minimum level of human error in targeting.”).

19. Schmitt & Thurnher, *supra* note 16, at 234 (“[N]o [autonomous] weapons have even left the drawing board.”).

20. LOSING HUMANITY, *supra* note 8, at 1-4; Kenneth Anderson, *Autonomous Weapon Systems and Their Regulation - A Flurry of Activity*, VOLOKH (Dec. 12, 2012, 9:32 PM), <http://www.volokh.com/2012/12/12/autonomous-weapon-systems-and-their-regulation-a-flurry-of-activity>; Noel Sharkey, *America’s Mindless Killer Robots Must Be Stopped*, GUARDIAN (Dec. 3, 2012, 11:00 AM), <http://www.guardian.co.uk/commentisfree/2012/dec/03/mindless-killer-robots>.

21. Schmitt & Thurnher, *supra* note 16, at 234.

be capable of adhering to the LOAC should be precluded, especially because such determinations often require a case-by-case analysis of factors.²² However, the speculative nature of predicting future technologies has not deterred HRW or its allies from initiating a witch-hunt against AWS.²³

There are portions of the LOAC that may be interpreted as providing a basis for banning a non-existent weapon.²⁴ The Martens Clause, when viewed as elevating the principles of humanity and the dictates of the public conscience to the level of independent sources of international law, would not require an analysis of the weapon's effects or the circumstances of its use in the same manner as would the traditional principles of the LOAC.²⁵ Under this view of the Martens Clause, it would be sufficient to proclaim a new weapon illegal under international law if the principles of humanity or the dictates of the public conscience were violated.²⁶ It is therefore the objective of this Note to explore the meaning of the Martens Clause, to address its implications on the burgeoning efforts to research and develop AWS, and to examine whether the Clause can be used to impose a preemptive prohibition on such weapons.

Part II will provide background on AWS and semi-autonomous systems, their existing precursors, and a brief overview of the possible evolution of related technologies.²⁷ Part II.A will define the terms used in this Note as well as those used in the reports, directives, and articles used in support.²⁸ Part II.A will also make explicit the difference between autonomous and semi-autonomous weapon systems, as well as in-the-loop, on-the-loop, and out-of-the-loop distinctions, with regard to where the human operator stands in relation to the weapon system.²⁹ Part II.B will establish a concise history of military robotics, and discuss certain semi-autonomous and automatic weapon systems that can be

22. See, e.g., *id.* at 254 (explaining that whether a principle of LOAC has been violated is "the product of a case-by-case assessment that is evaluated in terms of its reasonableness given the attendant circumstances").

23. See LOSING HUMANITY, *supra* note 8, at 46-47; *About Us*, CAMPAIGN TO STOP KILLER ROBOTS, <http://www.stopkillerrobots.org/about-us> (last visited July 18, 2013).

24. See Antonio Cassese, *The Martens Clause: Half a Loaf or Simply Pie in the Sky?*, 11 EUR. J. INT'L L. 187, 188-92 (2000).

25. See LOSING HUMANITY, *supra* note 8, at 30-36.

26. *Id.* at 25-26 ("[E]ven if a means of war does not violate an existing treaty or customary law, it can still be found unlawful if it contravenes the principles of humanity or the dictates of public conscience.").

27. See *infra* Part II.

28. See *infra* Part II.A.

29. See *infra* Part II.A.

viewed as forerunners of truly autonomous weapons.³⁰ Part II.C will discuss the future of AWS as outlined in official policy documents of the United States, and will explore the anticipated benefits of such weapons in terms of military and humanitarian usefulness.³¹ Part II.C will additionally cover the technological advances that will be necessary to achieve true autonomy in weapons.³²

The focus of Part III is to examine the arguments supporting a preemptive ban of AWS through the lens of international law, and to compare HRW's urged ban of AWS with existing prohibitions and restrictions of other weapons.³³ Part III.A will analyze *Losing Humanity*, a recent and prominent effort to prohibit AWS before they can be developed.³⁴ Part III.B will briefly explain relevant portions of the LOAC and its sources, and provide a historical overview of the Martens Clause.³⁵ Part III.C will discuss the prohibition of blinding laser weapons and other arms control agreements, and will analyze the role that the Martens Clause played in their accomplishment.³⁶ Lastly, Part III.D will scrutinize the Martens Clause in the context of a preemptive prohibition on the use and development of AWS.³⁷

Part IV will explore the likelihood of the Martens Clause contributing to a prohibition on AWS, as well as the positions that interested states should take in opposing such a prohibition.³⁸ Part IV.A will reexamine the interpretations of the Clause, and classify the Clause into three "benchmark" categories for evaluation.³⁹ After finding that a broad interpretation of the Clause poses the most significant and imminent risk to AWS, Part IV.B–C will survey and suggest various steps that interested states should take in order to protect themselves against opponents of AWS.⁴⁰ Additionally, Part IV.D will address the unprecedented nature of banning a weapon that has not yet been invented, let alone fielded in combat.⁴¹

Finally, Part V will draw conclusions about the impact the Martens Clause may have on AWS, the likelihood of its success in forging a

30. See *infra* Part II.B.

31. See *infra* Part II.C.

32. See *infra* Part II.C.

33. See *infra* Part III.

34. See *infra* Part III.A.

35. See *infra* Part III.B.

36. See *infra* Part III.C.

37. See *infra* Part III.D.

38. See *infra* Part IV.

39. See *infra* Part IV.A.

40. See *infra* Part IV.B–C.

41. See *infra* Part IV.D.

prohibition, and the decisions that interested states will have to navigate in the near future.⁴²

II. THE BACKGROUND AND TECHNOLOGIES OF AUTONOMOUS WEAPON SYSTEMS

'[A] robot must obey the orders given it by human beings except where such orders would conflict with the First Law.'

– Isaac Asimov, *The Second Law of Robotics*⁴³

A. Definitions

For the purpose of this Note, a weapon is defined as a device, implement, munition, substance, object, or piece of equipment that is capable of directing an offensive force toward a combatant or military objective.⁴⁴ The means of warfare, accordingly, refers to “all weapons, weapons platforms, and associated equipment used directly to deliver force during hostilities.”⁴⁵ Lastly, the term “[m]ethods of warfare” is defined as the application of such weapons in an armed conflict.⁴⁶ Robots are defined herein as machines that operate with some degree of autonomy and have the ability to sense and act in an environment according to their programming.⁴⁷ In the context of military application in armed conflicts, this Note regards robots as weapons or weapon systems, not soldiers or other actors in hostilities.

In addition, for the purposes of this discussion, the distinction between tiers of autonomy should be clarified. HRW and IHRC, in the recent Report calling for the prohibition of “killer robots” that autonomously target and kill,⁴⁸ define the levels of lethal autonomy as “Human-in-the-loop Weapons,” “Human-on-the-loop Weapons,” and “Human-out-of-the-loop Weapons.”⁴⁹ Human-in-the-loop Weapons, as described by the Report, are “[r]obots that can select targets and deliver force only with a human command”⁵⁰ This definition denotes semi-

42. See *infra* Part V.

43. I, ROBOT, *supra* note 1, at 37.

44. WILLIAM H. BOOTHBY, *WEAPONS AND THE LAW OF ARMED CONFLICT* 4 (2009).

45. *Id.*

46. *Id.*

47. LOSING HUMANITY, *supra* note 8, at 2.

48. *Id.* at 1. The HRW Report has been described as “a quasi-brief intended to justify a sweeping call for a preemptive, prohibitory multilateral treaty that would ban the ‘development, production, and use’ of autonomous weapons systems.” Anderson, *supra* note 20 (quoting LOSING HUMANITY, *supra* note 8, at 5).

49. LOSING HUMANITY, *supra* note 8, at 2.

50. *Id.*

autonomous capabilities in that the robot may perform some tasks without human intervention but still requires a human operator to acquire a target or strike.⁵¹ The Report defines Human-*on-the-Loop* Weapons as “[r]obots that can select targets and deliver force under the oversight of a human operator who can override the robots’ actions”⁵² Such a robot would be autonomous, as it does not require human interference or assistance to act.⁵³ Lastly, Human-*out-of-the-Loop* Weapons are defined by the Report as “[r]obots that are capable of selecting targets and delivering force without any human input or interaction.”⁵⁴ Such robots could be considered the most autonomous, as they are capable of acting—indeed, killing—entirely on their own.⁵⁵ These classifications, carefully chosen by the Report, are more concerned with the implementation of military robots and where the human operator sits within or upon the chain of command, rather than the attributes or capabilities of the machine itself.⁵⁶

Although the Report defines Human-*in-the-loop* Weapons as semi-autonomous, being unable to “select targets and deliver force” without human intervention,⁵⁷ this definition excludes a sizeable gray area in which a semi-autonomous weapon may select, acquire, or affirm the target autonomously but not strike autonomously, or, in the reverse, where the target may be preselected by a human operator and the weapon determines or implements the force that is issued.⁵⁸ The DoD, in its Directive addressing autonomy in weapons systems, defined the term “semi-autonomous weapon system” as a system which, “once activated, is intended to only engage individual targets or specific target groups that have been selected by a human operator.”⁵⁹ Such systems include:

Semi-autonomous weapon systems that employ autonomy for engagement-related functions including, but not limited to, acquiring, tracking, and identifying potential targets; cueing potential targets to human operators; prioritizing selected targets; timing of when to fire; or providing terminal guidance to home in on selected targets,

51. *See id.*

52. *Id.* The DoD, in the Directive, refers to such systems as “human-supervised autonomous weapon systems.” DIRECTIVE NO. 3000.09, *supra* note 14, at 13-14.

53. *See* LOSING HUMANITY, *supra* note 8, at 2.

54. *Id.*

55. *See id.*

56. *See id.*

57. *Id.*

58. *See id.*; DIRECTIVE NO. 3000.09, *supra* note 14, at 14.

59. DIRECTIVE NO. 3000.09, *supra* note 14, at 14.

provided that human control is retained over the decision to select individual targets and specific target groups for engagement.⁶⁰

Within this Note, the lethality of a robotic weapon system is defined as semi-autonomous when the weapon—otherwise capable of autonomously carrying out its task—requires a human operator for either target acquisition or target engagement, or both.⁶¹

The Report, in defining Human-*on-the-loop* and Human-*out-of-the-loop* Weapons alike as fully autonomous, is generally in accord with the DoD definition of an AWS, despite the shortcomings of the Report's definitions.⁶² It should be noted, however, that the DoD has stated that fully autonomous weapon systems have, at this time, only been authorized "for local defense" of manned installations and platforms, and are operated under the supervision of a human operator (Human-*on-the-loop*).⁶³ Furthermore, these AWS may only select non-human targets.⁶⁴ For the purpose of this Note, a fully autonomous weapon system is that which can select, acquire, or affirm its target and engage this target without human intervention.⁶⁵

B. A Brief History of Modern Military Robotics

Humans have always engaged in armed conflict, and this reality has often led to heated competition between states for supremacy and invention in "the arts and the instruments of force."⁶⁶ Presently, robotic weapon systems are also employed with levels of automation such that the human operator can assume a posture of observation, rather than control every action of the robotic system.⁶⁷ Examples of highly automated, human-supervised weapon systems include the Phalanx anti-

60. *Id.*

61. *See id.*

62. *See* LOSING HUMANITY, *supra* note 8, at 2; DIRECTIVE NO. 3000.09, *supra* note 14, at 14. The DoD defines an autonomous weapon system as a system that, "once activated, can select and engage targets without further intervention by a human operator," whether a human operator retains a supervisory override capability or not. DIRECTIVE NO. 3000.09, *supra* note 14, at 13-14.

63. DIRECTIVE NO. 3000.09, *supra* note 14, at 3.

64. *Id.* However, non-lethal autonomous weapons that pass the rigorous testing outlined in the Directive may select and engage targets for non-kinetic, i.e., electronic, attacks. *Id.*

65. *See id.* at 13-14.

66. KENNETH N. WALTZ, THEORY OF INTERNATIONAL POLITICS 127 (1979) ("The fate of each state depends on its responses to what other states do. The possibility that conflict will be conducted by force leads to competition in the arts and the instruments of force. . . . Contending states imitate the military innovations contrived by the country of greatest capability and ingenuity."); Jack M. Beard, *Law and War in the Virtual Era*, 103 AM. J. INT'L L. 409, 411 (2009); *see also* Aaron Saenz, *War 2.0 – Rise of the Robots*, SINGULARITY HUB (June 1, 2009, 12:17 PM), <http://www.singularityhub.com/2009/06/01/war-20-rise-of-the-robots> [hereinafter Saenz, *War 2.0*].

67. Saenz, *War 2.0*, *supra* note 66.

ship defense system, the Patriot anti-aircraft missile batteries, the Counter-Rocket, Artillery and Missile ("C-RAM") system, and the SGR-A1 sentry robot.⁶⁸

The Phalanx system provides anti-ship missile defense to various naval vessels by fully automatically conducting the functions of searching, detecting, threat evaluating, target acquisition, tracking, and firing.⁶⁹ The Phalanx automatically determines whether an approaching target is a threat or a non-threat by comparing the trajectory of the target against the software's threat logic.⁷⁰

On land, the C-RAM provides defense in hostile territories through automated turret intervention against surprise ballistic attacks.⁷¹ The C-RAM detects, tracks, targets, and eliminates incoming ballistics "faster than a human could even begin to command it to fire."⁷² The human's role in the operation of the C-RAM system is relegated to that of maintenance: the operator merely turns on and off the turret.⁷³ This human-machine relationship, which relies almost entirely on the robot, is by design: The C-RAM is meant to respond to threats that human beings simply are not fast enough to handle.⁷⁴

The SGR-A1 is a semi-autonomous sentry robot that scans the demilitarized zone ("DMZ") dividing North Korea and South Korea, and which is capable of identifying targets, issuing commands to surrender, observing signs of surrender, and responding appropriately.⁷⁵ The SGR-A1 can also attack its target without direct human control when it has been set to operate automatically.⁷⁶ It is important to note that the SGR-A1 merely identifies *every* target that enters the DMZ as an enemy by virtue of the person or object having entered the prohibited zone.⁷⁷

68. See Gary E. Marchant et al., *International Governance of Autonomous Military Robots*, 12 COLUM. SCI. & TECH. L. REV. 272, 276 (2011); Saenz, *War 2.0*, *supra* note 66; Jean Kumagai, *A Robotic Sentry for Korea's Demilitarized Zone*, IEEE SPECTRUM (Mar. 2007), <http://www.spectrum.ieee.org/robotics/military-robots/a-robotic-sentry-for-koreas-demilitarized-zone>.

69. M.S. Frick, *RAM and Phalanx: System of Systems Testing*, NAVY LEAGUE, <http://www.navyleague.org/seapower/sept00/frick.htm> (last visited July 18, 2013).

70. Robert H. Stoner, *R2D2 with Attitude: The Story of the Phalanx Close-In Weapons System (CIWS)*, NAVWEAPS, http://www.navweaps.com/index_tech/tech-103.htm (last updated Oct. 30, 2009).

71. Saenz, *War 2.0*, *supra* note 66.

72. *Id.*

73. *Id.*

74. *See id.*

75. *Samsung Techwin SGR-A1 Sentry Guard Robot*, GLOBALSECURITY.ORG, <http://www.globalsecurity.org/military/world/rok/sgr-a1.htm> (last visited July 18, 2013).

76. *Id.*

77. *Id.*

These semi-autonomous and automated weapon systems, when viewed as forerunners of AWS, provide examples of certain beneficial purposes for which AWS may be used in the future: such systems will be useful, if not crucial, in roles where speed is necessitated, unwavering persistence is required, and human soldiers simply cannot get the job done.⁷⁸

C. *The Future of Autonomous Weapon Systems*

Though the technology necessary to achieve full autonomy in weapon systems is not yet viable, the benefits of such systems has already become an area of interest in many military and robotics communities, and the United States has indicated its interest in developing such weapons in various military policy documents.⁷⁹

1. U.S. Policy Documents Regarding Autonomy in Future Weapon Systems

Military policy documents and directives issued by the United States have already laid the groundwork for increased deployment of unmanned systems and reflect plans to gradually increase the autonomy of weapon systems.⁸⁰ Some military and robotics experts have predicted that the technology required to establish truly autonomous weapons could be available within a few decades.⁸¹ Werner J.A. Dahm, Chief Scientist of the U.S. Air Force ("USAF"), stated in a 2010 report that, "by 2030 machine capabilities will have increased to the point that humans will have become the weakest component in a wide array of systems and processes."⁸² Recently, the United States became the first nation to release an official policy statement on the development of AWS—the Directive—in which the protocols for designing, testing, and fielding such systems are outlined.⁸³

78. See *supra* text accompanying notes 69–74; see also P.W. SINGER, *WIRED FOR WAR: THE ROBOTICS REVOLUTION AND CONFLICT IN THE TWENTY-FIRST CENTURY* 63–64 (2009). Additionally, robots are ideal candidates to undertake tasks that are dirty, dull, and dangerous. SINGER, *supra*, at 63.

79. See *infra* Part II.C.1–3.

80. LOSING HUMANITY, *supra* note 8, at 7.

81. *Id.* at 8.

82. U.S. AIR FORCE CHIEF SCIENTIST (AF/ST), REPORT ON TECHNOLOGY HORIZONS: A VISION FOR AIR FORCE SCIENCE & TECHNOLOGY DURING 2010–2030, at x (2010), available at <http://www.af.mil/shared/media/document/AFD-100727-053.pdf>.

83. See DIRECTIVE NO. 3000.09, *supra* note 14, at 1–4 (outlining U.S. policy for the development and implementation of AWS); see also Mark Gubrud, *DoD Directive on Autonomy in Weapon Systems*, INT'L COMM. FOR ROBOT ARMS CONTROL (Nov. 27, 2012), <http://www.icrac.net/2012/11/dod-directive-on-autonomy-in-weapon-systems>.

The USAF Unmanned Aircraft Systems Flight Plan (“Flight Plan”) envisions “a family of unmanned aircraft,” drastically ranging in size and function, all of which will include autonomous capabilities.⁸⁴ The USAF seeks to “harness increasingly automated, modular, globally connected, and sustainable multi-mission unmanned systems”⁸⁵ In the short-term, the USAF plans to incorporate autonomy for functions such as takeoff and landing, flight, swarming, and carrying out non-lethal mission operations.⁸⁶

The Flight Plan acknowledges that “[a]dvances in computing speeds and capacity will change how technology affects the [Observe, Orient, Decide, and Act] loop.”⁸⁷ As a result, “humans will no longer be ‘in the loop’ but rather ‘on the loop’ – monitoring the execution of certain decisions. Simultaneously, advances in [artificial intelligence] will enable systems to make combat decisions and act within legal and policy constraints without necessarily requiring human input.”⁸⁸ Noting that the implementation of lethal AWS in the battlefield “is contingent upon political and military . . . resol[ution of] legal and ethical questions,”⁸⁹ the Flight Plan’s long-term vision culminates with fully autonomous technology.⁹⁰ The weapon systems envisioned may be capable of autonomous target engagement, resulting in “a revolution in the roles of humans in air warfare.”⁹¹ Similar modernization programs are underway in the U.S. Army.⁹²

84. U.S. AIR FORCE, UNMANNED AIRCRAFT SYSTEMS FLIGHT PLAN 2009-2047, at 3 (2009), available at http://www.fas.org/irp/program/collect/uas_2009.pdf.

85. *Id.*

86. *Id.* at 3, 33. It is anticipated that swarming weapon systems will be capable of synchronized flight and attack, either autonomously or semi-autonomously. *See id.* at 16.

87. *Id.* at 41. The Observe, Orient, Decide, and Act (“OODA”) loop is a continuous process by which a machine observes and analyzes information, makes decisions, and acts upon such decisions. *Id.* By 2047, “technology will be able to reduce the time to complete the OODA loop to micro or nanoseconds” *Id.*

88. *Id.*

89. *Id.*

90. *Id.* at 50.

91. *Id.*

92. ARMIN KRISHNAN, KILLER ROBOTS: LEGALITY AND ETHICALITY OF AUTONOMOUS WEAPONS 71 (2009).

2. The Benefits of Autonomous Weapons

It has been posited by military and robotics experts that AWS may be an improvement over human soldiers in a variety of areas crucial to both military effectiveness and humanitarian concerns.⁹³ Ronald C. Arkin has stated that, “[i]n the fog of war it is hard enough for a human to be able to effectively discriminate whether or not a target is legitimate. Fortunately, it may be anticipated . . . that in the future autonomous robots may be able to perform better than humans under these conditions”⁹⁴

Some of the ethical advantages AWS will have over human soldiers may include: the ability to act more conservatively than their human counterparts due to the autonomous system’s lack of motivation for self-preservation;⁹⁵ the programming to behave in a self-sacrificing manner, if necessary, due to the absence of fear of death;⁹⁶ the ability to act without emotions, making autonomous systems insusceptible to anger or fear on the battlefield and therefore able to exercise clearer judgment than humans;⁹⁷ and immunity from psychological “scenario fulfillment,” which occurs “where humans use new incoming information in ways that only fit their pre-existing belief patterns,” and can result in “distortion or neglect of contradictory information in stressful situations.”⁹⁸

Arkin cites a report by the Mental Health Advisory Team (“MHAT”) that assesses the ethics and mental health of soldiers who had been deployed in Iraq and notes that the effects of battle on their psyches are alarming.⁹⁹ The MHAT report found that “[t]he level of combat is the main determinant of a Soldier’s or Marine’s mental health status” and that “[d]eployment length was related to higher rates of mental health problems”¹⁰⁰ Additionally, AWS may present numerous operational benefits to the military, including: “Faster, cheaper, better mission accomplishment; longer range, greater persistence, longer endurance, higher precision; faster target

93. See, e.g., ARKIN, *supra* note 6, at 29-31.

94. *Id.* at 29.

95. *Id.*

96. *Id.*

97. *Id.* at 29-30.

98. *Id.* at 30.

99. *Id.* at 31-32.

100. MENTAL HEALTH ADVISORY TEAM (MHAT) IV, FINAL REPORT: OPERATION IRAQI FREEDOM 05-07, at 3 (2006), available at <https://s3.amazonaws.com/s3.documentcloud.org/documents/551721/mental-health-advisory-team-mhat-iv.pdf>. This report also found that soldiers who experienced high levels of combat were nearly twice as likely to mistreat non-combatants as those who had experienced low levels of combat. *Id.* at 4.

engagement; and immunity to chemical and biological weapons among others.”¹⁰¹ Dr. Paul Kaminski, the Chairman of the Defense Science Board, succinctly stated the purpose of pursuing AWS: “The true value of these systems is not to provide a direct human replacement, but rather to extend and complement human capability in a number of ways.”¹⁰²

3. Technologies Required for True Autonomy

Critics and proponents of AWS debate whether technology will evolve to such a degree that fully autonomous weapon systems may be viable, effective, and compliant with the LOAC and international human rights concerns.¹⁰³ Despite the disjointed views about what the capabilities of currently non-existent future technologies will be, it is apparent that the “incremental march” towards the development and use of AWS is already underway.¹⁰⁴ Furthermore, serious questions loom over the fate of artificial intelligence (“AI”), the future state of which has been “overpromised” in the past.¹⁰⁵ Critics of AWS also assert that, aside from LOAC constraints, machine programming will never “replace the key elements of human emotion and affect that make human beings irreplaceable in making lethal decisions on the battlefield”¹⁰⁶ While those who oppose a preemptive ban of AWS admit that AI may never progress to such a level as to meet the requirements of the LOAC, it is hardly an established fact that human emotion is an irreplaceable, let alone safer, trait for decision-making on the battlefield.¹⁰⁷

101. ARKIN, *supra* note 6, at 30 (citation omitted).

102. DEF. SCI. BD., DEP’T OF DEF., TASK FORCE REPORT: THE ROLE OF AUTONOMY IN DoD SYSTEMS 1 (2012), available at <http://www.fas.org/irp/agency/dod/dsb/autonomy.pdf>.

103. Kenneth Anderson & Matthew Waxman, *Human Rights Watch Report on Killer Robots, and Our Critique*, LAWFARE (Nov. 26, 2012, 8:33 AM), <http://www.lawfareblog.com/2012/11/human-rights-watch-report-on-killer-robots-and-our-critique> [hereinafter Anderson & Waxman, *Human Rights Watch Report*].

104. Kenneth Anderson & Matthew Waxman, *Law and Ethics for Autonomous Weapon Systems: Why a Ban Won’t Work and How the Laws of War Can*, HOOVER INSTITUTION, Apr. 9, 2013, at 4-6, available at <http://www.hoover.org/publications/monographs/144241> [hereinafter Anderson & Waxman, *Law and Ethics*]. In fact, Kenneth Anderson and Matthew Waxman conclude that “[t]he incremental . . . deployment of autonomous weapon systems is inevitable” *Id.* at 27.

105. *Id.* at 14; cf. Schmitt, *Autonomous Weapon Systems*, *supra* note 9, at 17 (“After all, artificial intelligence is artificial.”).

106. Anderson & Waxman, *Law and Ethics*, *supra* note 104, at 14-15.

107. *Id.* (explaining that AWS, compared to human soldiers, may “reduce risks to civilians” in the future); see also *supra* text accompanying notes 93-104.

III. THE ARGUMENT FOR BANNING AUTONOMOUS WEAPON SYSTEMS, A REVIEW OF PAST WEAPONS PROHIBITIONS AND RESTRICTIONS, AND THE MARTENS CLAUSE

'[A] robot must protect its own existence as long as such protection does not conflict with the First or Second Laws.'

– Isaac Asimov, *The Third Law of Robotics*¹⁰⁸

New technologies have continuously produced increasingly complex means and methods of doing combat,¹⁰⁹ and there is little historical precedent for prohibiting new weapons before they have been fielded, let alone invented—yet, that is precisely what *Losing Humanity* seeks to accomplish.¹¹⁰

A. *The Debate Against Autonomous Weapons*

A recent debate has emerged surrounding AWS, and the critics have decried the inevitable failure of AWS to comply with LOAC principles.¹¹¹ These critics, most notably HRW, urge the preemptive ban on related research, development, and testing of lethal autonomous systems.¹¹² Most recently, the United Nations has joined the debate over AWS, concluding in an annual report (“U.N. Report”) that autonomous weapons should be approached with “great caution” and recommending

108. I, ROBOT, *supra* note 1, at 37.

109. Beard, *supra* note 66, at 411.

110. LOSING HUMANITY, *supra* note 8, at 5; Michael N. Schmitt, *Bellum Americanum: The U.S. View of Twenty-First Century War and Its Possible Implications for the Law of Armed Conflict*, 19 MICH. J. INT'L L. 1051, 1051-1052 (1998) [hereinafter Schmitt, *Bellum Americanum*]. The only weapons in modern history to have been banned before they were fielded in battle are blinding laser weapons, which were prohibited after they were viable but before they were implemented in combat. Protocol IV on Blinding Laser Weapons to the Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to Be Excessively Injurious or to Have Indiscriminate Effects, *opened for signature* Oct. 13, 1995, 2024 U.N.T.S. 167 [hereinafter Protocol IV]; Schmitt, *Bellum Americanum*, *supra*, at 1051-52.

111. See generally LOSING HUMANITY, *supra* note 8 (arguing that AWS will be unable to comply with LOAC tenets of distinction, proportion, military necessity, and the Martens Clause).

112. See LOSING HUMANITY, *supra* note 8, at 5; Tom Malinowski, Op-Ed., *A Dangerous Future of Killer Robots*, WASH. POST, Nov. 23, 2012, at A23; *Berlin Statement*, INT'L COMM. FOR ROBOT ARMS CONTROL, <http://www.icrac.net/statements> (last visited July 18, 2013) (urging the prohibition of “[f]urther development, acquisition, deployment, and use of armed autonomous robot weapons”); Sharkey, *supra* note 20 (“It is clear that the rational approach to the inhumanity of automating death by machine is to prohibit it.”). In April of 2013, HRW spearheaded the Campaign to Stop Killer Robots, a coalition of international NGOs established to further its crusade against AWS. *About Us*, *supra* note 23. The Campaign to Stop Killer Robots was recently described by Greg McNeal as a “fear campaign” that uses “scare tactics to simplify and amplify messages . . .” Greg McNeal, *Fear and the Killer Robots Meme*, FORBES (Feb. 27, 2013, 6:59 PM), <http://www.forbes.com/sites/gregorymcneal/2013/02/27/fear-and-the-killer-robots-meme>.

a moratorium on the development of AWS.¹¹³ The U.N. Report, unlike *Losing Humanity*, seeks to establish an international body “to monitor the situation and articulate the options for the longer term,” rather than demanding a preemptive prohibition of AWS outright.¹¹⁴

Losing Humanity urges all nations to “[p]rohibit the development, production, and use of fully-autonomous weapons through an international legally binding instrument” and to “[a]dopt national laws and policies to prohibit the development, production, and use of fully autonomous weapons.”¹¹⁵ The HRW and IHRC Report does not only avail itself of states, but also tasks roboticists and specialists with “[e]stablish[ing] a professional code of conduct governing the research and development of autonomous robotic weapons . . . to ensure that legal and ethical concerns about their use in armed conflict are adequately considered”¹¹⁶

On its most basic level, the Report espouses a “Frankensteinian fantasy”¹¹⁷ that “killer robots” will be uncontrollable, kill civilians, and, in the hands of a “repressive dictator,” be used to hunt down innocent citizens.¹¹⁸ The Report claims that AWS, when viable, will fail to comply with the LOAC.¹¹⁹ The Report alleges that the future technologies that will enable AWS will fail to meet the requirements of distinction, proportionality, military necessity, and the Martens Clause.¹²⁰ While there is some overlap between these principles, the focus of this Note is on the Martens Clause and its twin pillars.

113. Special Rapporteur on Extrajudicial, Summary or Arbitrary Executions, *Report on Lethal Autonomous Robots*, Human Rights Council, ¶¶ 1, 109, 113, U.N. Doc. A/HRC/23/47 (Apr. 9, 2013) (by Christof Heyns) [hereinafter U.N. Report].

114. Compare *id.* at ¶ 112 (recommending the establishment of an international body to oversee the development of AWS), with *LOSING HUMANITY*, *supra* note 8, at 5 (recommending that all nations “[p]rohibit the development, production, and use of fully autonomous weapons through an international legally binding instrument”).

115. *LOSING HUMANITY*, *supra* note 8, at 5.

116. *Id.*

117. Brad Allenby, *What Human Rights Watch's "Case Against Killer Robots" Gets Wrong About Military Realities*, SLATE (Nov. 20, 2012, 3:32 PM), http://www.slate.com/blogs/future_tense/2012/11/20/human_rights_watch_s_case_against_killer_robots_report_misunderstands_the.html (“The fundamental problem with [HRW’s] report . . . is its adoption of the Frankenstein worldview. The ‘humans vs. technology’ myth is quite powerful, especially in Western cultures, and always popular with audiences. But it is just too flawed and oversimplistic a foundation on which to build policy formulation.”); see also McNeal, *supra* note 112 (identifying the rhetoric of HRW and its Report as a “fear campaign” aimed at rousing public horror).

118. See *LOSING HUMANITY*, *supra* note 8, at 3-4, 46.

119. *Id.* at 30-36.

120. *Id.*

B. *The Law of Armed Conflict and the Martens Clause*

The LOAC is part of the broader body of public international law,¹²¹ the primary sources of which are customary law and multinational treaties.¹²² Other sources of authority specifically impacting Weapons Law are arms control agreements; although not always sources of international law, and not necessarily governed by international law, arms control agreements may also provide a legal basis for restricting or prohibiting the use of certain weapons.¹²³ Customary international law arises from the consenting conduct and practices of nations who engage in armed conflicts, with the caveat that the nation-actors must believe that their conduct and practices arise from legal requirements.¹²⁴ When determining the behavior and beliefs of nations for the purpose of establishing customary norms, numerous material sources of varying weightiness are consulted.¹²⁵ A treaty is “an international agreement concluded between States in written form and governed by international law, whether embodied in a single instrument or in two or more related instruments and whatever its particular designation”¹²⁶ Treaties are generally only binding upon the nations who are parties to the document and have consented to its terms as written.¹²⁷

The Martens Clause was originally recorded in the 1899 Hague Convention with Respect to the Laws and Customs of War on Land (“Hague Convention II”).¹²⁸ The Clause was then reformulated and

121. BOOTHBY, *supra* note 44, at 22.

122. *Id.* at 23.

123. *See id.* at 23, 29-30. Arms control agreements are agreed to by, and enforceable against, participating states, and are generally created to prohibit or restrict specific weapons or munitions research, production, design, or use. *See id.*; *see generally* Thomas Axworthy & Ryan Dean, *A Scan of Existing Arms Control Agreements with Lessons Learned*, INTERACTION COUNCIL (May 29-31, 2011), <http://www.interactioncouncil.org/scan-existing-arms-control-treaties-lessons-learned> (reviewing various arms control agreements).

124. BOOTHBY, *supra* note 44, at 23-25.

125. *Id.* at 27. These sources may include official policy statements and press releases, the opinions of legal advisors, legal manuals including manuals of military law and orders to armed forces, executive decisions and practices, international and national judicial decisions, treaties ratified by the state in question, and the practices of international organs. *Id.*

126. Vienna Convention on the Law of Treaties, art. 2, *opened for signature* May 23, 1969, 1155 U.N.T.S. 331.

127. BOOTHBY, *supra* note 44, at 28. Treaty obligations can at times give rise to customary norms that exist separately and concurrently to the treaty provisions themselves. *See id.* at 27.

128. *Opened for signature* July 29, 1899, 32 Stat. 1803, 1805. The Martens Clause, as originally recorded, stated:

Until a more complete code of the laws of war is issued, the High Contracting Parties think it right to declare that in cases not included in the Regulations adopted by them, populations and belligerents remain under the protection and empire of the

included in the 1907 Preamble to Convention IV with Respect to the Laws and Customs of War on Land ("Hague Convention IV").¹²⁹ The Hague Convention IV formulation of the Martens Clause states:

Until a more complete code of the laws of war has been issued, the High Contracting Parties deem it expedient to declare that, in cases not included in the Regulations adopted by them, the inhabitants and the belligerents remain under the protection and the rule of the principles of the law of nations, as they result from the usages established among civilized peoples, from the laws of humanity, and the dictates of the public conscience.¹³⁰

The Martens Clause was originally intended to preserve the customary rules that were in effect before the codification of the Hague Conventions and which were not named in those treaties.¹³¹ The precise current meaning of the Martens Clause and its impact on the LOAC is debated, and its interpretations are varied.¹³² The narrowest reading of the Martens Clause stands for the proposition that pre-existing customary law norms are not proscribed by positive treaty provisions.¹³³ The broadest interpretation of the Clause is that actors in an armed

principles of international law, as they result from the usages established between civilized nations, from the laws of humanity, and the requirements of the public conscience

Id. at 1805.

129. *Opened for signature* Oct. 18, 1907, 36 Stat. 2277, 2279-80 [hereinafter Hague Convention IV].

130. *Id.* The Martens Clause was subsequently incorporated into Additional Protocol I to the Geneva Conventions. Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of International Armed Conflicts, art. 1, *opened for signature* Dec. 12, 1977, 1125 U.N.T.S. 3, 7 [hereinafter Additional Protocol I]. The Clause, as expressed in Additional Protocol I, states that, "[i]n cases not covered by this Protocol or by other international agreements, civilians and combatants remain under the protection and authority of the principles of international law derived from established custom, from the principles of humanity and from the dictates of public conscience." *Id.*

131. Theodor Meron, *The Martens Clause, Principles of Humanity, and Dictates of Public Conscience*, 94 AM. J. INT'L L. 78, 79-80 (2000). Meron also explains that the modernized version of the Clause, recorded in Additional Protocol I, serves a similar purpose:

The Geneva Conventions employ a version of the Martens clause . . . for a somewhat different, but parallel, goal: to make clear that if they denounce the Conventions, the parties will remain bound by the principles of the law of nations, as they result from the usages established among civilized peoples, the laws of humanity, and the dictates of public conscience. This provision thus guarantees that international customary law will still apply for states no longer bound by the Geneva Conventions as treaty law.

Id. at 80.

132. Rupert Ticehurst, *The Martens Clause and the Laws of Armed Conflict*, 317 INT'L REV. RED CROSS 125, 126 (1997).

133. *Id.*

conflict are bound not only by customary norms and treaty law, but also by principles of humanity and the dictates of the public conscience.¹³⁴

At the very least, in the years since its introduction, the Clause has been frequently cited¹³⁵ and generally understood as extending to the entirety of the LOAC.¹³⁶ In *Legality of the Threat or Use of Nuclear Weapons* (“*Nuclear Weapons*”),¹³⁷ the International Court of Justice (“ICJ”) observed that the Martens Clause, and the fundamental rules it incorporates, is “to be observed by all States whether or not they have ratified the conventions that contain them, because they constitute intransgressible principles of international customary law.”¹³⁸

1. The Principles of Humanity and the Dictates of the Public Conscience

In *Nuclear Weapons*, the ICJ found that the Martens Clause “has proved to be an effective means of addressing the rapid evolution of military technology”¹³⁹ by preserving customary principles of humanity embodied in the Martens Clause, including the principles of distinction and proportionality,¹⁴⁰ the prohibition against unnecessary suffering,¹⁴¹ and the premise that the means of warfare are not unlimited.¹⁴² The principles of humanity, however, are already norms of customary international law, codified in treaties, or both—therefore, the traditional

134. *Id.*

135. Meron, *supra* note 131, at 78-79. Meron notes that:

[T]he Martens clause has been relied upon in the Nuremberg jurisprudence, addressed by the International Court of Justice and human rights bodies, and reiterated in many humanitarian law treaties that regulate the means and methods of warfare. It was restated in the 1949 Geneva Conventions for the Protection of Victims of War, the 1977 Additional Protocols to those Conventions, and the Preamble to the Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons, albeit in slightly different versions. The Martens clause . . . is cited or otherwise referred to in several national military manuals, including those of the United States, the United Kingdom, and Germany.

Id. at 78 (footnotes omitted).

136. *Id.* at 79.

137. Advisory Opinion, 1996 I.C.J. 226 (July 8).

138. *Id.* at 257.

139. *Id.* The ICJ’s statement on the effectiveness of the Clause actually does little to clarify its exact meaning, and the judges of the ICJ themselves could not agree on a singular interpretation of the Clause. See Ticehurst, *supra* note 132, at 127-28.

140. Schmitt, *Autonomous Weapon Systems*, *supra* note 9, at 14-15, 18. Distinction and the prohibition against weapons which are absolutely incapable of discrimination are different, but related, principles. *Id.* at 10.

141. *Id.* at 8-9.

142. Additional Protocol I, *supra* note 129, at 21.

interpretation of this prong of the Clause does not represent any original threat to AWS.¹⁴³

There is no singular meaning of the dictates of the public conscience.¹⁴⁴ Theodor Meron explains that this phrase can be examined from two distinct perspectives: first, “as public opinion that shapes the conduct of the parties to a conflict and promotes the development of international humanitarian law, including customary law”; and second, “as a reflection of *opinio juris*.”¹⁴⁵ Meron distinguishes the two perspectives in that public opinion has the power to promote the growth of customary law by persuading a government to regard “certain developing norms as already declaratory of customary law,”¹⁴⁶ whereas *opinio juris* is the belief of states and their authoritative figures that a particular practice is obligatory, or accepted as law—a predicate to the formation of customary norms.¹⁴⁷ Although public opinion and *opinio juris* are separate concepts, they are linked in that popular opinion often influences and induces government opinion.¹⁴⁸ Furthermore, Dan Belz has observed that public opinion also imposes “audience costs” on states as punishment for violations of international humanitarian law.¹⁴⁹ Therefore, nations may be incentivized to comply with humanitarian law in order “to win the support of public opinion,” avoid backlash in domestic elections, and avoid aggravating the civilian population who reside where the fighting occurs.¹⁵⁰

143. *Legality of the Threat or Use of Nuclear Weapons*, 1996 I.C.J. at 257; see, e.g., Additional Protocol I, *supra* note 129, at 21-22 (prohibiting the use of “weapons . . . and methods of warfare of a nature to cause superfluous injury or unnecessary suffering” and mandating the distinction between military targets and nontargets). The principles of humanity have, however, “been invoked rhetorically in attempts to humanize the behavior of parties using certain methods of warfare.” Meron, *supra* note 131, at 84. Furthermore, some states have argued that the principles of humanity prong of the Clause could “transform” general principles of humanity into “prohibitions on conduct,” without such principles having “ascended” to the status of customary norms. Michelle A. Hansen, *Preventing the Emasculation of Warfare: Halting the Expansion of Human Rights Law into Armed Conflict*, MIL. L. REV., Winter 2007, at 1, 19. It should also be noted that, in regards to the principles of humanity, some critics of AWS have suggested that “[t]aking humans out of the loop also risks taking humanity out of the loop,” but this author finds such arguments too literal to roundly address. U.N. Report, *supra* note 113, at ¶ 89.

144. See Hansen, *supra* note 143, at 19-20; cf. Meron, *supra* note 131, at 83-85.

145. Meron, *supra* note 131, at 83.

146. *Id.*

147. *Id.*

148. *Id.*

149. Dan Belz, *Is International Humanitarian Law Lapsing into Irrelevance in the War on International Terror?*, 7 THEORETICAL INQUIRIES L. 97, 121 (2006) (internal quotation marks omitted).

150. *Id.*

2. Modern Interpretations of the Martens Clause

The modern meaning of the Martens Clause is the subject of debate amongst judges, scholars, states, and NGOs.¹⁵¹ Antonio Cassese identifies three interpretations of the Martens Clause yielded from existing scholarship: the Clause helps to interpret existing principles and rules of international law; the Clause has elevated the principles of humanity and the dictates of the public conscience to the status of independent sources of international law; or the Clause has simply influenced and motivated the development of international law principles.¹⁵² The Statute of the International Court of Justice¹⁵³ recognizes another principle, albeit without referring to the dictates of the public conscience: “The Court, whose function is to decide in accordance with international law such disputes as are submitted to it, shall apply . . . the general principles of law recognized by civilized nations”¹⁵⁴ This interpretation offers a “public” composed only of states—civilized ones, at that—and therefore presents a relatively conservative view of the Clause.¹⁵⁵

A nuanced variation of Cassese’s second interpretation of the Clause urges that the principles of humanity and the dictates of the public conscience are independent sources of international law, but would require the exact content of these sources to be determined by courts of law in light of evolving circumstances.¹⁵⁶ Such a view would not immediately transform the principles of humanity and dictates of the public conscience into hard legal standards, but would permit courts to examine the acts of states (though, not necessarily acts which rise to the level of customary norms) as an additional basis for crystallizing these principles into international law.¹⁵⁷

The United States has subscribed to the view that the Martens Clause merely clarifies the existence and applicability of customary international law not explicitly addressed by the agreements contained in

151. See *LOSING HUMANITY*, *supra* note 8, at 26; Hansen, *supra* note 143, at 19-20. The debate over the meaning of the Clause is illustrated by the varying opinions issued in *Nuclear Weapons. Compare Legality of the Threat or Use of Nuclear Weapons*, Advisory Opinion, 1996 I.C.J. 226, 260 (July 8) (finding that the Martens Clause applies only “as an affirmation that the principles and rules of humanitarian law apply to [newly created weapons]”), with *id.* at 408-11 (Shahabuddeen, J., dissenting) (arguing that the Martens Clause is not “confined to supplying a humanitarian standard by which to interpret separately existing rules of conventional or customary international law”).

152. Cassese, *supra* note 24, at 189-92; accord Ticehurst, *supra* note 132, at 126.

153. June 26, 1945, 59 Stat. 1031, 1055.

154. *Id.* at 1060.

155. See Cassese, *supra* note 24, at 189-92.

156. *Id.* at 191.

157. *Id.* at 191-92.

multinational conventions.¹⁵⁸ In practice, this interpretation of the Martens Clause is the most widely accepted, as “no domestic or international court has fleshed out the scope of these dictates as independent sources of law, and the Martens Clause has never been successfully invoked to preclude the use of a new weapon.”¹⁵⁹ In fact, guided by such an interpretation, the Martens Clause alone would never be sufficient to prohibit a new weapon.¹⁶⁰

On the other hand, there are humanitarian NGOs that espouse the broadest view of the Martens Clause.¹⁶¹ The International Committee of the Red Cross (“ICRC”) has stated that “[a] weapon which is not covered by existing rules of international humanitarian law would be considered contrary to the Martens [C]lause if it is determined *per se* to contravene the principles of humanity or the dictates of public conscience.”¹⁶² HRW, maintaining a similar position, has noted that the LOAC is shaped not only by reference to existing customary and treaty law, but also by reference to the principles of humanity and the dictates of the public conscience.¹⁶³

The overwhelming majority of cases that have interpreted the Martens Clause have used it in a supplemental fashion.¹⁶⁴ The Clause has thus been generally cited to advance the idea that the principles of humanity and dictates of the public conscience are supplemental sources of international law used to bolster a finding that a practice contravenes international law on another, more traditional or substantial legal principle.¹⁶⁵ In such instances, therefore, the Martens Clause was inessential because decisions were reached on other grounds.¹⁶⁶ These supplemental, *ad abundantiam* uses of the Clause “primarily . . . pa[id]

158. See Hansen, *supra* note 143, at 19.

159. Lesley Wexler, *Limiting the Precautionary Principle: Weapons Regulation in the Face of Scientific Uncertainty*, 39 U.C. DAVIS L. REV. 459, 483 (2006); see Cassese, *supra* note 24, at 208.

160. See Cassese, *supra* note 24, at 189-92. If the Clause’s dual prongs are not elevated to the status of independent sources of international law, the Clause would necessarily have to attach to a traditional source of law—such as a customary norm or treaty provision—to effectively invalidate a new weapon. See *id.*

161. See, e.g., Int’l Comm. of the Red Cross, *A Guide to the Legal Review of New Weapons, Means and Methods of Warfare: Measures to Implement Article 36 of Additional Protocol I of 1977*, 88 INT’L REV. RED CROSS 931, 945 (2006), available at http://www.icrc.org/eng/assets/files/other/irrc_864_icrc_geneva.pdf (adopting the broadest view of the Martens Clause).

162. *Id.*

163. HUMAN RIGHTS WATCH, *BLINDING LASER WEAPONS: THE NEED TO BAN A CRUEL AND INHUMANE WEAPON* 31 (1995), available at <http://www.hrw.org/reports/1995/General1.htm> (“[T]he Martens clause establishes that despite continuing developments in technology and methods of warfare, the principles of custom, humanity, and public conscience always hold.”).

164. Cassese, *supra* note 24, at 202.

165. *Id.* at 202-07.

166. *Id.*

lip service to humanitarian demands, rather than . . . support[] the notion that” principles of humanity and dictates of the public conscience are independent, singly effective sources of international law.¹⁶⁷ The persuasiveness of this view is buttressed by the fact that no international or national court has ever held that a rule emerged from the principles of humanity or the dictates of the public conscience, unlike other, more traditional pillars of the LOAC.¹⁶⁸ Accordingly, a fair, but certainly not the only, interpretation could view the Clause as establishing a framework for interpreting international rules, rather than creating them.¹⁶⁹ That is, courts can and should consider the principles of humanity as well as the dictates of the public conscience, which, while not determinative, may serve as guidelines for evaluating the issue before them.¹⁷⁰

C. *Past Weapons Prohibitions and the Martens Clause*

In recent history, certain weapons have been prohibited or restricted through the creation of multinational treaties.¹⁷¹ One of the most prominent of which is the Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May be Deemed to be Excessively Injurious or to Have Indiscriminate Effects (“CCW”)¹⁷² and its annexed protocols.¹⁷³ The weapons prohibited by the CCW are those that may be deemed to be excessively injurious or to have injurious effects, rather than those that *have been deemed* to have such effects, evincing a careful avoidance of any determination that the weapons addressed in the treaty were *per se* illegal before the treaty was put into effect.¹⁷⁴ This raises the question of whether the prohibitions and

167. *Id.* at 208.

168. *Id.*

169. *Id.*

170. *Id.*

171. See, e.g., Protocol for the Prohibition of the Use in War of Asphyxiating, Poisonous or Other Gases, and of Bacteriological Methods of Warfare, *opened for signature* June 17, 1925, 26 U.S.T. 571 [hereinafter Geneva Gas Protocol] (prohibiting the use of gas and poisonous weapons in armed conflicts).

172. *Opened for signature* Apr. 10, 1981, 1342 U.N.T.S. 137 [hereinafter Convention on Conventional Weapons].

173. See, e.g., Protocol IV, *supra* note 110, at 167.

174. BOOTHBY, *supra* note 44, at 115-16; W. Hayes Parks, *Means and Methods of Warfare*, 38 GEO. WASH. INT’L L. REV. 511, 518 (2006) [hereinafter Parks, *Means and Methods*]. Parks notes that:

[H]ad the weapons in question been deemed by governments to be ‘excessively injurious or to have indiscriminate effects’ *per se* . . . their use already would be prohibited, making new treaty provisions unnecessary. This indicates—subsequently confirmed by conference results—that the conference and convention title were more political rhetoric

restrictions created under the CCW are actually based upon the LOAC, or if they are merely arms control agreements.¹⁷⁵

Of particular interest to this Note are the following three weapons prohibitions: the Geneva Convention Protocol on the Prohibition of the Use in War of Asphyxiating, Poisonous or Other Gases, and of Bacteriological Methods of Warfare (“Geneva Gas Protocol”);¹⁷⁶ the CCW’s first Annexed Protocol on Non-Detectable Fragments;¹⁷⁷ and the CCW’s fourth Annexed Protocol on Blinding Laser Weapons (“Protocol IV”).¹⁷⁸ The Geneva Gas Protocol is of interest because the language contained therein specifically contemplates the “general opinion of the civilised world,”¹⁷⁹ whereas the CCW’s first and fourth Protocols each yielded the prohibition of non-existent weapons.¹⁸⁰

The Geneva Gas Protocol states:

WHEREAS the use in war of asphyxiating, poisonous or other gases, and of all analogous liquids, materials or devices, has been *justly condemned by the general opinion of the civilised world*; and

....

TO THE END that this prohibition shall be universally accepted as a part of International Law, binding alike the *conscience* and the practice of nations;

....

[T]he High Contracting Parties . . . accept this prohibition . . .¹⁸¹

While the Geneva Gas Protocol is noted here for its inclusion of notions of general opinion, as well as its mention of the conscience of nations—implicating, if indirectly, the Martens Clause—the Geneva Gas Protocol is essentially a *reactive* treaty addressing the effects of gas

than an accurate picture of the views of the majority of governments.

Id.

175. *Id.* at 519.

176. *Opened for signature* June 17, 1925, 26 U.S.T. 571.

177. Convention on Conventional Weapons, *supra* note 169, at 168.

178. Protocol IV, *supra* note 110, at 167.

179. Geneva Gas Protocol, *supra* note 173, at 575.

180. Parks, *Means and Methods*, *supra* note 174, at 519-20, 526 (explaining that the Protocols prohibiting non-detectable fragments and blinding laser weapons banned non-existent weapons). Note, however, that there were claims that the United States was secretly developing laser weapons that would blind enemies. See Jack H. McCall, Jr., *Blinded by the Light: International Law and the Legality of Anti-Optic Laser Weapons*, 30 CORNELL INT’L L.J. 1, 6 (1997). *Contra* W. Hays Parks, *The ICRC Customary Law Study: A Preliminary Assessment*, 99 AM. SOC’Y INT’L L. PROC. 208, 211-12 (2005) (“The Customary Law Study statement that ‘Prior to the 1995 Protocol IV to the CCW, the [United States] was developing a number of laser systems intended to blind either personnel and/or optical systems,’ is categorically false.” (alteration in original)).

181. Geneva Gas Protocol, *supra* note 171, at 575 (emphases added).

warfare.¹⁸² Furthermore, the codification of the prohibition as a treaty indicates that such methods of warfare were not *already* illegal under the amorphous Martens Clause.¹⁸³

The Protocol on Non-Detectable Fragments precludes a non-existent weapon in a manner that is also largely directed at the weapon's effects; it states that "[i]t is prohibited to use any weapon *the primary effect* of which is to injure by fragments which in the human body escape detection by X-Rays."¹⁸⁴ Even if such weapons were to be created, they are not likely to be designed for the primary purpose of leaving non-detectable fragments inside enemies—rather, where non-detectable fragments, such as plastics, had been used in weapons before, the primary purpose was to make the weapons lighter or reduce production costs.¹⁸⁵ And, though this prohibition was *proactive* in that it was based on the determination that non-detectable fragments would cause superfluous injury, these effects were predictable: non-detectable fragments, by definition, cannot be detected, making them unnecessarily difficult to remove.¹⁸⁶

Similarly, Article 1 of Protocol IV states that:

It is prohibited to employ laser weapons specifically designed, as their sole combat function or as one of their combat functions, to cause permanent blindness to unenhanced vision, that is to the naked eye or to the eye with corrective eyesight devices. The High Contracting Parties shall not transfer such weapons to any State or non-State entity.¹⁸⁷

In prohibiting laser weapons to the extent that their sole or primary combat function was to cause permanent blindness, it appears that the principle of unnecessary suffering was the underlying force that drove the agreement.¹⁸⁸ However, Protocol IV most certainly acts upon the dictates of the public conscience, as well.¹⁸⁹ Much of the discourse

182. *See id.*

183. *See supra* text accompanying note 174.

184. Convention on Conventional Weapons, *supra* note 172, at 168 (emphasis added).

185. Parks, *Means and Methods*, *supra* note 174, at 519-20. If the primary purpose served to make the weapons lighter or cheaper, and non-detectable fragments were left in the wounded incidentally, or even secondarily, such a weapon would not come under the prohibition of the protocol. *Id.*; *see* Convention on Conventional Weapons, *supra* note 172, at 168.

186. *See* Parks, *Means and Methods*, *supra* note 174, at 519-20; *see also* Convention on Conventional Weapons, *supra* note 172, at 168.

187. Protocol IV, *supra* note 110, at 167.

188. *See id.*

189. *See id.*; *see also* Additional Protocol I, *supra* note 130, at 21; Ann Peters, *Blinding Laser Weapons: New Limits on the Technology of Warfare*, 18 LOY. L.A. INT'L & COMP. L.J. 733, 752-53, 755-56 (1996). Note that the military utility deprived by the prohibition of permanently blinding

around the prohibition of blinding laser weapons focused on the fact that permanently blinding lasers were perceived as cruel, inhumane weapons; therefore, they would be contrary to the dictates of the public conscience—or, at least, public opinion would disfavor them.¹⁹⁰

In calling for the prohibition of blinding laser weapons, HRW noted that experts it had consulted:

[W]ere largely in agreement that laser weapons and methods of warfare that cause blindness would run counter to the requirements of established custom, humanity, and public conscience. Some experts expressed either personal repugnance for lasers or the belief that their countries' civilian population would find the use of blinding as a method of warfare horrific.¹⁹¹

HRW found that public opinion could be more negatively affected by blinding lasers than by weapons that kill outright due to the fact that the blinded veterans would remain in public view with their disabilities.¹⁹² HRW also found that blinding laser weapons could negatively impact peace negotiations and societal infrastructure by leaving so many wounded veterans alive, but permanently maimed.¹⁹³

Although blinding laser weapons had not been fielded by any belligerent in an armed conflict at the time that they were prohibited, the effects of blinding laser weapons were substantially known.¹⁹⁴ For instance, following World War I, stark “images of blinded, shuffling columns of soldiers, robbed of sight by the use of chemical agents such as mustard gas, left a lingering impression of the horrors of . . . warfare.”¹⁹⁵ The after-effects of permanent blinding were also widely known to include psychological injuries and a shorter life expectancy.¹⁹⁶ Since the effects of blinding laser weapons would

laser weapons is slight, considering that temporarily blinding laser weapons, not covered by the agreement, serve the same or similar purpose. Schmitt, *Autonomous Weapon Systems*, *supra* note 9, at 36 n.100.

190. See, e.g., HUMAN RIGHTS WATCH, *supra* note 163, at 22-32 (detailing the humanitarian and legal considerations for prohibiting permanently blinding laser weapons).

191. *Id.* at 31.

192. *Id.* at 32.

193. *Id.*; accord McCall, *supra* note 180, at 12-13. McCall stated:

Permanent blindness is quite unlike any other type of wound inflicted in warfare. Most battle casualties—approximately sixty percent—recover fully from their wounds. . . . Unlike most combat wounds, however, laser blinding would likely create a large category of permanently disabled survivors, potentially taxing the resources of any nation's medical and socio-economic systems.

Id.

194. See McCall, *supra* note 180, at 43.

195. *Id.*

196. *Id.*

predictably cause many of the same—if not identical—results as blinding chemical weapons, the “public outcry and revulsion” that was caused by blinding gases was substantially analogous to that of blinding laser weapons.¹⁹⁷ As such, the prohibition of poisonous, chemical weapons served as a blueprint for the prohibition of blinding laser weapons, and mitigated the fact that there was little or no empirical data regarding the effects of the blinding laser weapons, which were never fielded combat.¹⁹⁸

D. The Martens Clause as the Basis for Prohibition of Autonomous Weapon Systems

HRW seeks a prohibition of the research, design, and use of AWS through the implementation of multinational treaties and national regulations.¹⁹⁹ Additionally, HRW argues that the Martens Clause provisions of principles of humanity and the dictates of the public conscience may also invalidate AWS before they come to exist.²⁰⁰ The Report finds that, “even if a means of war does not violate an existing treaty or customary law, it can still be found unlawful if it contravenes the principles of humanity or the dictates of public conscience.”²⁰¹ Accordingly, the Report interprets the Clause as providing either supplemental weight to arguments based on traditional pillars of LOAC, or as a fallback argument in case a multinational prohibitory treaty should not gain much traction.²⁰² The Report states that the Clause:

which encompasses rules beyond those found in treaties, requires that means of warfare be evaluated according to the “principles of humanity” and the “dictates of public conscience.” Both experts and laypeople have an [sic] expressed a range of strong opinions about whether or not fully autonomous machines should be given the power to deliver lethal force without human supervision. While there is no consensus, there is certainly a large number for whom the idea is shocking and unacceptable. States should take their perspective into account when determining the dictates of public conscience.²⁰³

The Report further asserts that “any review of fully autonomous weapons should recognize that for many people these weapons are

197. *Id.*

198. *Id.*; see *supra* text accompanying notes 192-95.

199. LOSING HUMANITY, *supra* note 8, 5.

200. *Id.* at 25-26.

201. *Id.* at 26.

202. See *id.*

203. *Id.* at 35 (citation omitted).

unacceptable under the principles laid out in the Martens Clause,” and also found that “fully autonomous weapons would likely contravene the Martens Clause, which prohibits weapons that run counter to the ‘dictates of public conscience.’”²⁰⁴ Embedded in HRW’s Martens Clause analysis is both an insistence that the Clause creates independently viable sources of international law—capable of deeming AWS *per se* illegal absent codified treaty provisions to the same effect—and the more generally accepted view that the Martens Clause may supplement more established treaty or customary principles of the LOAC.²⁰⁵

IV. OPPOSING A PREEMPTIVE BAN ON AUTONOMOUS WEAPON SYSTEMS

‘A robot may not injure humanity or, through inaction, allow humanity to come to harm.’

– Isaac Asimov, *The Zeroth Law of Robotics*²⁰⁶

A. *Categorizing the Martens Clause Interpretations*

The interpretation of the Martens Clause, and the weight afforded to the principles of humanity and the dictates of the public conscience, will determine how the Clause can impact—or prevent—the development or use of AWS in armed conflicts.²⁰⁷ As discussed in Part III, the Martens Clause has been interpreted in a variety of ways.²⁰⁸ For the purpose of brevity, three benchmark interpretations of the Clause—one Narrow, one Moderate, and one Broad—will herein be identified and analyzed for their present potential to prohibit AWS.²⁰⁹

The Narrow View interprets the Martens Clause as merely ensuring that customary norms of international law are neither preempted by treaty provisions, nor nullified by the absence of customary norms from positive treaty provisions.²¹⁰ The Narrow View does not deem the principles of humanity or the dictates of the public conscience to be independent sources of international law.²¹¹ Under this interpretation,

204. *Id.* at 4, 36.

205. *See id.* at 4, 26, 35.

206. FOUNDATION AND EARTH, *supra* note 2, at 485. The Zeroth Law was conceived by the robots themselves to supersede the First Law of Robotics—which dictates that a robot may not harm a human being—and permits the possibility of harming individual human beings in order to preserve humanity as a whole. *See id.*

207. *See discussion supra* Part III.

208. *See supra* Part III.B.

209. *See infra* text accompanying notes 210-22.

210. Ticehurst, *supra* note 132, at 126.

211. *See id.*

states alone would have the final authority to declare AWS illegal before they exist, either through the creation of treaties or the establishment of customary norms.²¹² As a result, the Martens Clause—standing for the proposition that customary international law is not precluded by the presence or absence of treaty provisions—would only affect the legality of AWS insofar as customary norms and treaty provisions *already* address the same.²¹³ Therefore, under the Narrow View, the Martens Clause alone could not be made to prohibit AWS.

The Moderate View of the Martens Clause identifies the principles of humanity and the dictates of the public conscience as supplemental sources of international law, capable of being used in conjunction with other principles of the LOAC to add additional support to the proposition that an action violates international law.²¹⁴ Under the Moderate View, the two prongs of the Martens Clause—the principles of humanity and the dictates of the public conscience—are not independently operable sources of law, as are treaty provisions and customary norms.²¹⁵ Relying on this interpretation, the Clause could *influence or strengthen* a determination that a non-existent weapon violates the LOAC, but the Clause alone would be insufficient to prohibit it.²¹⁶ Under the Moderate View—requiring that the principles of humanity or the dictates of the public conscience attach to a treaty provision or customary rule of international law—the Clause is not presently a serious threat to AWS because the effect of existing treaties and customs on AWS is debated and unresolved.²¹⁷ Furthermore, the principles of the LOAC relied on in *Losing Humanity* largely require a case-by-case weighing of factors related to the non-existent weapon's capabilities, use, and effects—none of which can be evaluated until AWS exist.²¹⁸

The Broad View of the Martens Clause regards the principles of humanity and the dictates of the public conscience as independently

212. See discussion *supra* Part III.B.

213. See Schmitt & Thurnher, *supra* note 16, at 275. Schmitt and Thurnher state plainly that the Martens Clause, “[b]y its own terms . . . applies only in the absence of treaty law. In other words, it is a failsafe mechanism meant to address lacunae in the law; it does not act as an overarching principle that must be considered in every case.” *Id.* (footnote omitted).

214. See Cassese, *supra* note 24, at 202-07. For example, the Clause’s principles of humanity and the dictates of the public conscience have been used to strengthen and support independently sufficient violations of treaty provisions. See *id.*

215. *Id.* at 202.

216. See *id.*; see also Schmitt, *Autonomous Weapon Systems*, *supra* note 9, at 32.

217. See Schmitt, *Autonomous Weapon Systems*, *supra* note 9, at 32.

218. See *id.* at 16; see also *LOSING HUMANITY*, *supra* note 8, at 30-35.

enforceable sources of international law.²¹⁹ Under this interpretation, belligerent actors in an armed conflict are bound not only by customary and treaty law, but also by the principles of humanity and the dictates of the public conscience.²²⁰ With these principles elevated to independent sources of international law, either principle could, if violated, theoretically prohibit a non-existent weapon.²²¹ The Broad View is NGO-friendly in that it may not require states to establish, for example, a new multinational treaty to invalidate a weapon that violates the principles of humanity or the dictates of the public conscience.²²² Therefore, the Broad View is the greatest threat to AWS at the present and, naturally, it is the interpretation that *Losing Humanity* propounds.²²³

B. *Advocating Against the Broad Interpretation*

Under the Narrow and Moderate Views of the Martens Clause, it is unlikely that AWS would face much, if any, threat of being preemptively prohibited because the Clause, so interpreted, does not elevate the principles of humanity or the dictates of the public conscience to independent sources of international law.²²⁴ Accordingly, it is imperative that states seeking to protect their interests in autonomous weapons object fiercely to interpretations of the Martens Clause that purport to enlarge the principles of humanity and the dictates of the public conscience.²²⁵ If NGOs are successful in their attempts to expand their own influence and law-making authority through expansive interpretations of the Martens Clause, the power of states to control their own means and methods of combat will be diminished.²²⁶ The principles of humanity and the dictates of the public conscience are well-established aids for interpreting and supplementing the traditional pillars of the LOAC, but “the Martens [C]ause does not allow one to build castles of sand. . . . [I]ts references to principles of humanity and dictates

219. Ticehurst, *supra* note 132, at 126.

220. *Id.*

221. *See id.*; *see also* Cassese, *supra* note 24, at 210-11 (explaining that some states have subscribed to the Broad View and regard the principles of humanity and the dictates of the public conscience as “legally binding standards”).

222. *See* Ticehurst, *supra* note 132, at 126.

223. *See* LOSING HUMANITY, *supra* note 8, at 25-26.

224. *See* discussion *supra* Part IV.A.

225. *See* discussion *supra* Parts II.A, IV.A.

226. *See* Meron, *supra* note 131, at 87-88 (“[The Martens Clause] serves as a powerful vehicle for governments and especially NGOs to push the law ever more to reflect human rights concerns. Where there already is some legal basis for adopting a more humanitarian position, the Martens [C]ause enables decision makers to take the extra step forward.”).

of public conscience cannot, alone, delegitimize weapons and methods of war, especially in contested cases.”²²⁷

To prevent the adoption of—or acquiescence to—the Broad View of the Clause, states must reject this interpretation.²²⁸ To do so, states may, for example, refuse to sign or participate in the formation of treaties that enlarge the scope of the Clause; domestically interpret the Clause narrowly in official documents, directives, and judicial decisions; directly speak out against enlarging interpretations of the Clause; and forge agreements or treaties with other states that adopt the Narrow or Moderate View.²²⁹ States can also engage each other to discuss the implications of an expanded Martens Clause, and expose the shortcomings of the Broad View.²³⁰ In particular, the difficulty of transforming the overly broad principles of humanity and the ever-changing dictates of the public conscience into independent sources of law should be reason enough for most states to prefer the Narrow or Moderate View of the Clause.²³¹

Also note that international actions, especially those in the realm of armed conflict, involve far more than just humanitarian concerns: the reality is that “power, reciprocity, and the interests of the parties involved” are major concerns, in spite of the growing influence of NGOs, the media, and the public.²³² While the Martens Clause—and the humanitarian interests it reflects—have achieved successes in restraint of the conduct of belligerents in armed conflicts, elevating the principles of humanity and the dictates of the public conscience will ultimately stretch the Clause thin and create an ambiguous field of moral law and public opinion that may undermine the traditional tenets of the LOAC.²³³ The Clause was simply not designed for such use.²³⁴

C. *State Countermeasures Under the Broad Interpretation*

The Martens Clause, broadly interpreted, is uniquely poised to prohibit AWS before the technology is developed or fielded in combat, should the principles of humanity or the dictates of the public conscience

227. *Id.* at 88.

228. *See id.*

229. *See id.*

230. *See id.* at 88-89.

231. *See id.*; *see also* Joseph Burns Kelly, *Gas Warfare in International Law*, 9 MIL. L. REV. 1, 62 (1960) (“The public conscience, though it may shape state policy, is too unreliable to bind states legally to a pattern of conduct for the future.”).

232. Meron, *supra* note 131, at 89.

233. *See id.* at 88.

234. *See supra* text accompanying notes 131-50.

so demand.²³⁵ This is true because the broad interpretation, favored by HRW, could empower NGOs to command the preemptive prohibition of AWS merely upon a showing of inhumanity or widespread public outcry, without relying upon the traditional principles of the LOAC.²³⁶ Naturally, the Broad View of the Martens Clause incentivizes the dissemination of sensationalist, fear-mongering rhetoric aimed at persuading the public, impressionable states, or NGOs that the challenged weapons are abhorrent and must be banned before they exist.²³⁷ Although the Broad View is not the generally accepted interpretation of the Clause, states should nonetheless be prepared to defend their interests against it.²³⁸ To do so, interested states should attempt to educate other nations, judiciaries, the public, and relevant organizations to influence the international marketplace of ideas, promote the potential benefits of AWS, and establish common standards of review and implementation for AWS.

1. Dispelling the Science Fiction Dystopia

States opposing a ban of AWS should make clear the distinction of science fiction and science fact, expose the irrationality of the fictional theme of robots turning on their creators, and explain that the development of AWS will follow the incremental growth of robotics technology.²³⁹ Directly addressing the principles of humanity and the dictates of the public conscience, *Losing Humanity* claims that, “[b]oth experts and laypeople have . . . expressed a range of strong opinions about whether or not fully autonomous machines should be given the power to deliver lethal force without human supervision. While there is no consensus, there is certainly a large number for whom the idea is shocking and unacceptable.”²⁴⁰ While the Report states that there is no consensus—therefore public opinion is currently undecided—the

235. See discussion *supra* Part IV.A–B.

236. See *LOSING HUMANITY*, *supra* note 8, at 35; Meron, *supra* note 122, at 88–89. The dictates of the public conscience may refer to the public at large, states themselves, public officials or semi-authoritative groups such as NGOs, depending on the meaning given to the term “public.” Meron, *supra* note 122, at 85.

237. See, e.g., *LOSING HUMANITY*, *supra* note 8, at 1 (sensationalizing “killer robots”); Kelly, *supra* note 227, at 61–62 (explaining that the public conscience, warped by propaganda, is unreliable); McNeal, *supra* note 112 (criticizing the “scare tactics” propagated by AWS’ critics).

238. Cassese, *supra* note 24, at 202 (noting that a majority of cases cited the Martens Clause in a supplemental manner); see *supra* Part IV.A–B.

239. See Anderson & Waxman, *Law and Ethics*, *supra* note 104, at 4, 11 (discussing the incremental development of autonomous weapons, and the benefit of building such weapons with principles of the LOAC in mind); see also *LOSING HUMANITY*, *supra* note 8, at 4 (confusing dystopian science fiction for science fact).

240. *LOSING HUMANITY*, *supra* note 8, at 35.

reasoning addresses the principles of humanity in an overly literal manner by equating Human-*out-of-the-Loop* systems with the Martens Clause's principles of humanity.²⁴¹ The Report, however, misses the point: a literally inhuman weapon system may prove to be far more humane than human soldiers could ever be.²⁴²

The Report's reasoning is actually propagandizing, and it is notably consistent with themes of science fiction entertainment²⁴³ as well as the sensationalist media coverage of AWS, and robots, generally.²⁴⁴ Itself no stranger to yellow journalism and eye-catching headlines, *Losing Humanity's* subtitle—*The Case Against Killer Robots*—shows that the publication connotes, and even lauds, the man versus machine myth.²⁴⁵

Science fiction has long been a successful forum for exploring the human condition and social issues, but it has also been a forum rich in horror, uncontrollable creations, and malevolence.²⁴⁶ The classic science fiction theme, referenced in *Losing Humanity*, is that our own creations will turn on us, resulting in "terrifying consequences."²⁴⁷ This is the theme of *Frankenstein*,²⁴⁸ as well as many modern movies,²⁴⁹ and "[t]he popular reaction . . . is a demand for more control immediately, whether or not it is advisable."²⁵⁰

241. See *id.* at 35-36.

242. See ARKIN, *supra* note 6, at 29-31; see also discussion *supra* Part II.C.2.

243. See, e.g., Christine Corcos et al., *Double-Take: A Second Look at Cloning*, *Science Fiction and Law*, 59 LA. L. REV. 1041, 1061-62 (1999) ("[Science fiction] films . . . depict future worlds destroyed by science and technology run amok . . ."); Malinowski, *supra* note 112, at A23 ("[T]he danger of delegating warfare to machines has been a central theme of modern science fiction.").

244. See, e.g., John Roach, *Ban Killer Robots Before They Go Berserk*, *Group Says*, NBC NEWS, <http://www.nbcnews.com/technology/futureoftech/ban-killer-robots-they-go-berserk-group-says-1C7182082> (last visited July 18, 2013) (reporting that an autonomous robot "could confuse a young girl with an ice cream cone for a soldier with a gun and kill her"); Sharkey, *supra* note 20 (referring to AWS as "mindless killer robots" and "autonomous killers"); see also Emma Hughes & Jenny Kitzinger, *Science Fiction Fears? An Analysis of How People Use Science Fiction in Discussing Risk and Emerging Science and Technology* 3 (Soc. Contexts & Responses to Risk Network, Working Paper No. 28, 2008) ("Film makers use technologies such as cloning or genetic research as themes in horror or science-fiction dystopias and these threatening themes or images are also borrowed by documentary makers and news journalists." (citation omitted)).

245. See generally LOSING HUMANITY, *supra* note 8 (rousing fear of "killer robots"); see also Allenby, *supra* note 107 ("The Frankensteinian label of 'killer robots' is good PR, and good fantasy noir, but blinks reality.").

246. Corcos et. al., *supra* note 243, at 1056.

247. *Id.*

248. See generally MARY SHELLEY, *FRANKENSTEIN* (Bantam ed. 2003) (depicting a tale of a man-made monster that turns on its creator).

249. Corcos et. al., *supra* note 243, at 1060-1063.

250. *Id.* at 1063.

To overcome the negative impact of dystopian futures appropriated by NGOs and the media, states seeking to develop and use AWS should educate nations, the public, and the media about the true nature of the technological art and the incremental development of autonomy that will enable such weapons.²⁵¹ Additionally, military secrecy in weapon development should be carefully navigated, and transparency should be strategically, but reasonably, employed.²⁵² The appearance of secrecy breeds distrust, as is made apparent in HRW's report accusing the United States of secretly testing blinding laser weapons in the decade before Protocol IV was enacted.²⁵³ Kenneth Anderson and Matthew Waxman suggest that the United States adopt a policy of transparency with regards to the testing and safety precautions—if not the programming or offensive capabilities—of AWS, thereby “shaping the normative terrain” and establishing global standards for AWS.²⁵⁴ By making the policy decisions surrounding AWS transparent, best practices can be established while technical results and processes can necessarily remain secret or shared only with allies to protect military advantage.²⁵⁵

Governments themselves may also issue official directives, policy statements, or legal reviews to establish their practices regarding AWS and to begin building a body of customary norms.²⁵⁶ The Directive—by establishing DoD policy, approval chains, review procedures, and safety precautions—addresses AWS responsibly and creates standards for their development, testing, and use that can be adopted as best practices by other states in the future.²⁵⁷ Furthermore, the Directive undercuts at least one of *Losing Humanity's* fears in that, absent special approval by several U.S. military authorities, fully autonomous weapons will not be permitted to target human beings.²⁵⁸

251. See Anderson & Waxman, *Law and Ethics*, *supra* note 104, at 24-25.

252. *Id.* at 25-26.

253. HUMAN RIGHTS WATCH, *supra* note 163, at 3-4.

254. Anderson & Waxman, *Law and Ethics*, *supra* note 104, at 3, 24-26.

255. *Id.* at 24-25.

256. See *id.* at 25-26; see also, e.g., DIRECTIVE NO. 3000.09, *supra* note 14, at 1 (establishing DoD policy “for the development and use of autonomous and semi-autonomous functions in weapon systems”).

257. See generally DIRECTIVE NO. 3000.09, *supra* note 14 (outlining the research, development, and use of AWS in the U.S. armed forces).

258. See *id.* at 3 (“Human-supervised autonomous weapon systems may be used to select and engage targets, with the exception of selecting humans as targets, for local defense to intercept attempted time-critical or saturation attacks for [defense of manned installations and platforms] . . .”).

2. Humanitarian Potential

States should reiterate what scholars, military experts, and roboticists have posited: in the future, AWS could be more compliant with the principles of the LOAC than human soldiers, and may permit armed conflicts to be fought with fewer inhumane consequences.²⁵⁹ These predictions not only help to allay fears of “killer robots” run amok, but also suggest that a prohibition of the research and development of AWS would not only be premature but irresponsible as well.²⁶⁰ Prohibiting the development of AWS “may have the effect of denying commanders a tool for minimizing the risk to civilians”²⁶¹

It is well documented that humans are imperfect soldiers, and the experience of war can cause some soldiers to violate the standards of the LOAC.²⁶² The effects of battle can increase a soldier’s propensity to commit atrocities and suffer mental breakdowns; studies have found that the more time a soldier spends in a battle zone, the more likely these results become.²⁶³ In this respect, AWS may have distinct advantages over human soldiers: machines experience no desire for self-preservation, anger, frustration, or fear of death.²⁶⁴

Losing Humanity ignores these facts and asserts that robots—lacking empathy and compassion—will cause more harm than human soldiers.²⁶⁵ HRW audaciously champions “the superiority of human emotions in controlling targeting and firing of weapons, [and] human empathy over human fear.”²⁶⁶ However, HRW’s “humanitarian” recommendation to preemptively ban AWS could actually result in a counter-humanitarian outcome: machines may “reduce risks to civilians by making targeting more precise and firing decisions more controlled

259. See ARKIN, *supra* note 6, at 29-30; Schmitt, *Autonomous Weapon Systems*, *supra* note 9, at 25; Don Troop, *Robots at War: Scholars Debate the Ethical Issues*, CHRON. HIGHER EDUC., Sept. 14, 2012, at A1; see also discussion *supra* Part II.C.2.

260. See Schmitt, *Autonomous Weapon Systems*, *supra* note 9, at 36-37.

261. *Id.* Schmitt also notes that “[i]t runs counter to the object and purpose of [the LOAC] to suggest that a weapon system that reduces harm to combatants in situations in which its use does not aggravate civilian risk should be unlawful.” *Id.* at 36.

262. ARKIN, *supra* note 6, at 29.

263. See *supra* text accompanying notes 99-100.

264. ARKIN, *supra* note 6, at 29-30.

265. *LOSING HUMANITY*, *supra* note 8, at 38. The Report suggests that AWS would be perfect for a repressive autocrat, because “[e]ven the most hardened troops can eventually turn on their leader if ordered to fire on their own people.” *Id.* The Report goes on to suggest that AWS, on the other hand, would crush the rebels because the “[r]obots would not identify with their victims and would have to follow orders no matter how inhumane they were.” *Id.*

266. Anderson, *supra* note 20; cf. *LOSING HUMANITY*, *supra* note 8, at 37-38.

[]especially compared to human-soldier failings that are so often exacerbated by fear, panic, vengeance, or other emotions”²⁶⁷

3. Establishing Common Standards for Testing and Evaluation

States who have an interest in the creation and use of AWS should foster a common set of standards for their evaluation and testing.²⁶⁸ Common standards would serve to legitimize efforts to create these weapons safely and in compliance with the LOAC from the ground up.²⁶⁹ Furthermore, common standards of review and compliance would prevent unscrupulous states from developing AWS with “few or no constraints at all.”²⁷⁰

D. An Unprecedented Prohibition

The prohibition on the use and transference of blinding laser weapons is a modern arms control agreement that banned a weapon that had never been fielded in combat.²⁷¹ There was, however, a surfeit of data on the effects of mass permanent blinding when Protocol IV was created.²⁷² The discourse surrounding this agreement was focused, in part, on the application of the Martens Clause and the implications of the public conscience.²⁷³ The resulting protocol, however, textually prohibits blinding laser weapons purely on the basis of the weapons’ effects.²⁷⁴

Attempts to prohibit specific weapons—particularly those that do not yet exist—on the basis of the dictates of the public conscience should be rejected. Before the dictates of the public conscience can be accepted, hypothetically, as an independent source of law, two questions must be definitively answered: What public? and Whose conscience?²⁷⁵ If these dictates are merely a matter of public opinion, then the Clause

267. Anderson & Waxman, *Law and Ethics*, *supra* note 104, at 15.

268. *Id.* at 24-26.

269. *See id.* at 25.

270. *Id.* at 3, 20.

271. Parks, *Means and Methods*, *supra* note 174, at 519-20, 526.

272. *See supra* text accompanying notes 186-98.

273. *See* McCall, *supra* note 180, at 29-32.

274. *See* Protocol IV, *supra* note 110, at 167 (“It is prohibited to employ laser weapons specifically designed . . . to cause permanent blindness” (emphasis added)). Likewise, the Protocol on Non-Detectable Fragments prohibits the use of certain weapons on the basis of those weapons’ effects. Convention on Conventional Weapons, *supra* note 172, at 168 (“It is prohibited to use any weapon the primary effect of which is to injure by [non-detectable] fragments” (emphasis added)). *But see* Geneva Gas Protocol, *supra* note 173, at 575 (prohibiting the use of certain gases in armed conflict on the basis that such gases had “been justly condemned by the general opinion of the civilised world”).

275. *Cf.* Meron, *supra* note 131, at 85.

would be overly vague and prone to endless fluctuations.²⁷⁶ If the public disagrees over what is conscionable, how will a judiciary decide? Perhaps more importantly, how will states know what practices are prohibited as result of the dictates of public conscience?

Even if the Martens Clause withstands attacks to its vague nature, AWS should not be banned on the tenuous predictions of NGOs or foreboding themes of science fiction.²⁷⁷ In the past, the High Contracting Parties to the Geneva Gas Protocol were spurred to act, in part, by the “horrific” nature of those weapons’ effects.²⁷⁸ Where the effects of such weapons were known, and were truly found to be abhorrent, the dictates of the public conscience were legitimately *considered*, although the Clause was never found to independently operate to prohibit a challenged weapon.²⁷⁹ Unlike permanently blinding lasers and poisonous gases, there is no history—devastating or otherwise—of the effects of AWS in combat: such systems do not yet exist.²⁸⁰ Therefore, claims that AWS will have horrifying, inhumane effects are presently just speculation.²⁸¹

V. CONCLUSION

‘How do you decide what is injurious, or not injurious, to humanity as a whole?’

– Isaac Asimov²⁸²

This Note concludes that, because AWS do not yet exist and there is no empirical data on the effects of such systems, the Martens Clause currently poses a discernible, but tenuous, threat to their legality.²⁸³ While there are many interpretations of the Clause, the Broad View applied by HRW could, if successful, establish the principles of humanity and dictates of the public conscience as independent standards of international law.²⁸⁴ *Losing Humanity* claims that the public

276. McCall, *supra* note 180, at 30-31.

277. See *supra* Part III.A–C.

278. McCall, *supra* note 180, at 31.

279. See discussion *supra* Part III.C.

280. See discussion *supra* Part III.C–D; see also *LOSING HUMANITY*, *supra* note 8, at 1.

281. See Anderson, *supra* note 20 (criticizing *Losing Humanity* as a “factually speculative report”).

282. FOUNDATION AND EARTH, *supra* note 2, at 486 (“In theory, the Zeroth Law was the answer to our problems. In practice, we could never decide. A human being is a concrete object. Injury to a person can be estimated and judged. Humanity is an abstraction. How do we deal with it?” (internal quotation marks omitted)).

283. See discussion *supra* Part IV.A–B.

284. See *LOSING HUMANITY*, *supra* note 8, at 35-36; see also *supra* Part IV.A–B.

conscience will—if it does not already—dictate that AWS are inhumane and abhorrent.²⁸⁵ Under the Broad View of the Clause, an authoritative finding that AWS violate the dictates of the public conscience could result in the prohibition of such weapons, whether they exist or not.²⁸⁶ Although this interpretation is by no means the majority view, it is not completely unsupported.²⁸⁷

Even if the Broad View becomes accepted, states can still protect their interests in developing autonomous weapons. In so doing, governments should issue policy statements and legal reviews of AWS that elucidate the scientific reality of such systems and dispel the dystopian fears disseminated by NGOs and the media.²⁸⁸ States should also emphasize the potential humanitarian benefits of AWS—to impact public opinion—and establish common standards for evaluating and testing AWS, creating a framework of best practices and customary norms.²⁸⁹ Finally, states should adamantly object to attempts to prohibit non-existent weapons to avoid setting adverse precedents in the area of weapons regulation.²⁹⁰ Though the future of autonomy in weapons systems is uncertain, a preemptive prohibition of such systems on the basis of the Martens Clause should not be forged.²⁹¹ Whether the Clause is a mere relic or something more remains to be seen, but its potential for expanded use in deciding matters of international law remains as present as ever, and “[a]t odd and unpredictable times, we cling in fright to the past.”²⁹²

Tyler D. Evans*

285. LOSING HUMANITY, *supra* note 8, at 35-36.

286. *See supra* Part IV.

287. *See* Ticehurst, *supra* note 123, at 128-29.

288. *See* discussion *supra* Part IV.C.1-2.

289. *See* discussion *supra* Part IV.C.1-3.

290. *See* discussion *supra* Part IV.C.

291. *See* discussion *supra* Part IV.

292. ISAAC ASIMOV, FOUNDATION'S EDGE 3 (Doubleday 1982).

* J.D. Candidate, 2014, Hofstra University School of Law; B.F.A., 2007, State University of New York at Purchase College. This Note is dedicated to my father, Paul W. Evans, and grandfather, Moe 'Red' Klang. I would foremost like to thank my mother, Jane Evans, for her endless support over the years, and Professor Michael Torlen for teaching me that the whole is greater than the sum of its parts. I am also indebted to Professor Kevin McElroy for his guidance, but not for his quips about the Mets; Brian Sullivan and Sarah Freeman for being a pleasure to work with; David Gerardi, Rebecca Sklar, and Stephen Piraino for showing us the ropes; Chris Rickard for the children's book about robots; Michal Ovadia for the vigilant research; Erik Harmon for the corrections in green ink; and James O'Connor for the blue. Finally, I would like to express my gratitude to Isaac Asimov for providing me with such eloquent and fascinating phrases to disperse throughout this Note. *See, e.g.,* ISAAC ASIMOV, FOUNDATION AND EMPIRE 103 (1st ed. 1952) (“It is the invariable lesson to humanity that distance in time, and in space as well, lends focus.”).
