LETHAL AUTONOMOUS WEAPONS AND THE END OF JUST WAR: AWAKENED AUTOMATA OR SOLEMN SIMULACRA?

BY

MAJOR CARLO ANGELO S. CORDERO

A THESIS PRESENTED TO THE FACULTY OF

THE SCHOOL OF ADVANCED AIR AND SPACE STUDIES

FOR COMPLETION OF GRADUATION REQUIREMENTS

SCHOOL OF ADVANCED AIR AND SPACE STUDIES

AIR UNIVERSITY

MAXWELL AIR FORCE BASE, ALABAMA

JUNE 2018

APPROVAL

The undersigned certify that this thesis meets master's-level standards of research, argumentation, and expression.

_____
STEPHEN D. CHIABOTTI (Date)

_____
THOMAS A. HUGHES (Date)

DISCLAIMER

The conclusions and opinions expressed in this document are those of the author. They do not reflect the official position of the US Government, Department of Defense, the United States Air Force, or Air University.

ABOUT THE AUTHOR

Major Carlo Angelo S. Cordero is a United States Air Force senior pilot with more than 2,200 flying hours in various aircraft. He earned his commission in 2005 through the Air Force ROTC program at the University of Maryland, College Park where he graduated with a Bachelor of Arts degree in Criminology/Criminal Justice. Major Cordero was selected for Joint Specialized Undergraduate Pilot Training; flying the T-6A Texan II, T-1 Jayhawk, and ultimately the KC-135R/T Stratotanker. After completing KC-135 initial training as a distinguished graduate, Major Cordero served as an Aircraft Commander, Standardization and Evaluation Liaison Officer, and Training Officer at the 91st Air Refueling Squadron, MacDill AFB, Florida. Later, he volunteered to cross-train into the Remotely Piloted Aircraft community and assigned to Creech AFB, Nevada. Here, Major Cordero became an Instructor Pilot in the MQ-1B Predator at the 18th Reconnaissance Squadron, Executive Officer at the 432d Operations Group, and Director of Staff/Pilot in the RQ-170 Sentinel program. Throughout his career, he has deployed numerous times to support overseas contingency operations. Major Cordero was a distinguished graduate of the Squadron Officer School, selected for Air Command and Staff College early, and chosen for advanced studies at the US Air Force's graduate school for strategists, the School of Advanced Air and Space Studies at Maxwell AFB, Alabama.

ACKNOWLEDGMENTS

My time at the School of Advanced Air and Space Studies has been challenging but has been one of the most rewarding experiences in my career. This is in no small part due to the expertise of the faculty who, through their continual guidance, have elevated class XXVII into remarkable thinkers and leaders. Explicitly, I give my thanks to my thesis advisor Dr. Stephen Chiabotti and reader Dr. Thomas Hughes, whose help throughout the process was priceless and whose discussions expanded my perspective. I am also very much appreciative of my classmates whose dialogue and debate throughout the year helped generate clarity of thought to my position.

I would also like to thank my friends and family - some who shared their passion for the subject matter and fueled my efforts, and others who, at my request, ventured outside their comfort zones to provide unbiased feedback to my task. Most importantly, I am especially grateful to my wife and kids who have shared in this experience every step of the way. I am truly indebted to my wife who kept the kids and I alive/sane through long days and nights and has always given me her unconditional support.

ABSTRACT

For warfighting soldiers, sailors, Airmen, and marines, the Just War Theory embodies much of their code of conduct that helps to limit the damage accepted as these good men and women take action to confront the evils of the world. Even if not directly applied by warfighters, the rules of engagement and war that govern conduct in battle, the decisions to resort to violence made by war-waging political and senior military leaders, and campaign objectives and lines of effort selected by operational commanders conform to the tenets of Just War Theory and so have an indirect, though still meaningful, effect.

The idea that, at some level, a responsible party can be held accountable for the morality of an action underpins Just War Theory. This liability may be moral or legal in nature. As technological advances allow humans to be farther from combat – both in physical distance and the point of decision – can this culpability still indeed be traced back to the operator, commander, programmer, or nation that employed the system? Alternatively, must the system itself be found at fault for a violation? If the machine can be held accountable, how does Just War Theory apply? If the machine cannot, how would Just War Theory restrict or permit its use?

Having the reputation as both the preeminent military might and global moral compass, the United States must be prepared to employ emergent technologies in a moral and just manner or, at least, be ready to respond to nations that do not. Questions of morality and Just War Theory applicability arise in both the employment and response to autonomous technology. Understanding the justice of the autonomous systems America's actions are normalizing as weapons is a critical component of the nation's war preparation. This paper will seek to answer the question: how does Just War Theory, specifically *jus in bello*, apply to lethal autonomous weapon systems?

Despite the notion that advanced weapon systems are independently choosing targets and can make life-and-death decisions, investigation reveals that this is not so. Examination of the points of decision and consequence show that moral obligation remains with the humans that choose to field these weapons and how they are employed. Even in the theoretical "independent" lethal weapon system scenario where a machine is sentient - envisioned by Hollywood and science fiction writers - moral agency cannot be attributed to these entities. As such, *jus in bello* principles apply to the users of these lethal weapon systems rather than to the systems themselves.

CONTENTS

Illustrations

Chapter 1

Who Cares?

For decades the United States has enjoyed uncontested or dominant superiority in every operating domain...Today, every domain is contested.

- James Mattis, United States Secretary of Defense

The next war will not decide who is right but who is left.

- Col. Frank P. Hobgood

It has often been said that "the only thing necessary for the triumph of evil is that good men do nothing."[1] Sometimes, however, good people must cause more harm before a confronted evil is stopped. Like the surgeon cutting into her patient to remove a malignant tumor, pain, blood, and risks are borne before there can be any healing. Doctors have long held to the Hippocratic Oath or some other code of ethics that guides their actions, aids in decisions, and minimizes acceptable harm to patients. For warfighting soldiers, sailors, Airmen, and marines, the Just War Theory embodies much of their code of conduct that helps to limit the damage accepted as these good men and women take action to confront the evils of the world. Even if not directly applied by warfighters, the rules of engagement and war that govern conduct in battle, the decisions to resort to violence made by war-waging political and senior military leaders, and campaign objectives and lines of effort selected by operational commanders, conform to the tenets of Just War Theory and so have an indirect, though still meaningful, effect.

The idea that, at some level, a responsible party can be held accountable for the morality of an action underpins Just War Theory. This liability may be moral or legal in nature. As technological advances allow humans to be farther from combat – both in physical distance and the point of decision – can this culpability still indeed be traced back to the operator, commander,

[1] "The Only Thing Necessary for the Triumph of Evil Is That Good Men Do Nothing," Quote Investigator, accessed April 10, 2018, https://quoteinvestigator.com/2010/12/04/good-men-do/.

programmer, or nation that employed the system? Alternatively, must the system itself be found at fault for a violation? If the machine can be held accountable, how does Just War Theory apply? If the machine cannot, how would Just War Theory restrict or permit its use?

In recent years, the utilization of Remotely Piloted Aircraft (RPA) in the Global War on Terror has garnered both praise and scorn from a variety of people, both domestic and international. Many decry the use of RPAs as immoral due to the significant asymmetric advantage given to the United States and its allies and the misperception that a machine is independently killing human beings. Other nations recognize the advantage and continue to develop their unmanned systems to challenge American asymmetry. As technological advances increase at an exponential rate, remotely piloted vehicles coupled with varying forms of Artificial Intelligence or, at a minimum, systems that can observe, orient, decide, and act faster than the humans that traditionally wage war, will be developed and fielded. Lethal autonomous weapon systems exist now to a certain degree and will only increase in lethality and ability as progress in technology multiplies the military advantage of these systems. It is no wonder that lethal autonomous weapon systems are heralded as "the third revolution in warfare, after gunpowder and nuclear arms."[2]

Since the 9/11 attacks on the United States, unmanned platforms like the MQ-1B Predator experienced startling growth in use and evolution. Its success on the battlefields of Afghanistan and Iraq highlighted the advantages of unmanned platforms and accelerated the development of more autonomous weapon systems. This growth has become a worldwide phenomenon encompassing government and commercial applications in all domains - air, land, and sea.[3] Secretary of Defense James Mattis, in the National Defense Strategy of 2018, warns of the dangers of rapid technological advancements changing the character of war due to a relentless drive for greater capabilities

---

[2] Amitai Etzioni, "Pros and Cons of Autonomous Weapons Systems (with Oren Etzioni)," Library of Public Policy and Public Administration Happiness Is the Wrong Metric, 2018, 256, doi:10.1007/978-3-319-69623-2_16.

[3] Jeffrey L. Caton, *Autonomous Weapon Systems: A Brief Survey of Developmental, Operational, Legal, and Ethical Issues* (Carlisle, PA: Strategic Studies Institute and U.S. Army War College Press, 2015), 62.

and a lower barrier of entry for actors.[4] The U.S. Government Accountability Office states in a July 2012 report that, in only seven years, the number of unmanned aerial vehicle (UAV)-equipped countries spiked from about 40 to more than 75.[5] The proliferation of the technology is a testament to the benefit garnered from these systems juxtaposed to the low cost of acquisition. This scenario applies to other types of autonomous technology as both the cyber and space battlefields continue to mature. The merits of autonomous systems may reveal them to be weapons of war *par excellence*.

There is a division, however, about the morality of the use of such systems. Many military experts believe the use of various autonomous weapon systems are morally preferable to the use of human combatants, while critics demand limitations on these weapons, or better yet, altogether banning them on the basis of moral and legal reasons.[6] Proponents call attention to advantages such as their force multiplicative effect allowing for fewer warfighters in a given engagement but with greater effect; expansion of the battlefield into previously inaccessible areas; reduction in human casualties by replacing them in high-danger missions; suitability for dull (long duration), dangerous (Explosive Ordinance Disposal), or dirty (Hazardous Materials) missions; potential lower-cost alternative to manpower (TALON combat robot costs $230,000 vs. the average per year cost of a soldier at $850,000); and heightened performance capabilities by rendering physiological concerns of pilots and soldiers irrelevant.[7] There is no denying the benefits of lethal autonomous weapon systems on the battlefield, but such advantage is inconsequential to those opposed. Their stance is best described by the 2014 North Atlantic Treaty Organization (NATO) Allied Command Transformation report which concludes, "the idea that autonomous systems could be autonomously deciding on the use of lethal force against humans, is perceived by some as being incompatible with the dictates of public conscience."[8]

---

[4] Mattis, *National Defense Strategy*, 3.
[5] Caton, *Autonomous Weapon Systems,* 8.
[6] Etzioni, Pros and Cons, 253.
[7] Etzioni, Pros and Cons, 253-254.
[8] Caton, *Autonomous Weapon Systems,* 45.

Secretary Mattis is adamant "the homeland is no longer a sanctuary," and so the nation must be prepared.[9] Doing so means that America and her allies must have the capability to orient, observe, decide, and act faster than the enemy. In addition to cultivating such ability in its warfighters, the Defense Department is investing and seeking advantages in cutting-edge autonomous systems to include military applications of autonomy, artificial intelligence, and machine learning.[10] The department's Third Offset Strategy and Project Maven initiative seek to exploit the advantages of technology to outmaneuver adversaries before the battle has even started.

Prussian general and noted military theorist Carl von Clausewitz writes that "war is an act of force…which theoretically can have no limits." He explains that the logic of war is one of escalation as "each of the adversaries forces the hand of the other," resulting in a "reciprocal action" driven by what both Athenian historian and general Thucydides and political philosopher Thomas Hobbes call a "necessity of nature."[11] This spiral into absolute and total war is not inevitable, for even Clausewitz asserts that the theoretical possibilities of war and the actual realities of war are not the same.[12] The existence of war at some level below absolute is simultaneously the cause and effect of Just War Theory. *Jus in bello* provides a framework from which just warriors can execute their wartime duties without impugning their honor or morality, and in a manner less likely to escalate the suffering already expected. In fact, "twentieth-century international law has attempted to recover the notion of just cause in war, thereby attempting to reassert the importance of *jus ad bellum*. Beginning with the League of Nations Covenant and the Pact of Paris (or Kellogg-Briand Pact) and continuing in articles 2 and 51 of the United Nations Charter, the effort has been to outlaw aggressive war while accepting defense as a legitimate cause for going to war."[13]

---

[9] Mattis, *National Defense Strategy*, 3.
[10] Mattis, *National Defense Strategy*, 7.
[11] Michael Walzer, Just and Unjust Wars a Moral Argument with Historical Illustrations (New York, NY: Basic Books, 2015), 23.
[12] Carl Von Clausewitz et al., *On War* (Princeton, NJ: Princeton University Press, 1989), 91-94.
[13] James Turner. Johnson, *Can Modern War Be Just?* (New Haven: Yale, 1986), 21.

The tenets of Just War Theory call on war-fighters and war-wagers to choose and act morally. Lethal autonomous weapon systems challenge the very notion of just fighting as some argue in favor of their moral applications and others contend the very notion is not possible. Roboticist Ronald Arkin believes "autonomous robots in the future will be able to act more 'humanely' on the battlefield for a number of reasons, including that they do not need to be programmed with a self-preservation instinct, potentially eliminating the need for a 'shoot first, ask questions later' attitude." Moreover, unclouded by emotions like fear and frenzy, lethal autonomous weapon systems have the potential to synthesize greater levels of information without external influence to generate the most accurate actionable assessments.[14] Still, it is not the battlefield prowess of the autonomous robot that is of concern, but "the delegation of life-or-death decision making to nonhuman agents."[15]

Having the reputation as both the preeminent military might and global moral compass, the United States must be prepared to employ emergent technologies in a moral and just manner or, at least, be ready to respond to nations that do not. Questions of morality and Just War Theory applicability arise in both the employment and response to autonomous technology. Mattis warns that the nation must be ready to act within a world of great-power competition. Understanding the justice of the autonomous systems America's actions are normalizing as weapons is a critical component of the nation's war preparation. This paper will seek to answer the question: how does Just War Theory, specifically *jus in bello*, apply to lethal autonomous weapon systems?

This chapter has introduced the central question of the paper and provides explanations of the importance of Just War Theory, lethal autonomous weapon systems, and the need to determine if just-war with autonomous systems is possible. This is primarily a paper about Just War Theory, set in the context of lethal weapon systems. The main thrust is not to explore the potential of future lethal weapon systems, but to evaluate the current technology with our understanding of Just War Theory.

---

[14] Etzioni, Pros and Cons, 255.
[15] Etzioni, Pros and Cons, 256.

Chapter 2 provides an essential exploration of Just War Theory. This initial examination will give readers a baseline understanding of the concepts of *jus in bello* that will be evaluated in the context of lethal weapon systems.

Chapter 3 focuses on the spectrum of autonomy applied to technology. The current use of the word "autonomy" has contributed to the confusion regarding the morality of autonomous weapon systems. Current definitions allow differentiation of systems within the Department of Defense and other organizations by focusing on the level of human interaction with the system, but this is problematic its implications because it encourages a lay audience to think machines are making life-and-death decisions without human input. This chapter focuses on a system's ability for mission completion as a stand-alone entity rather than its relationship with a human operator. This perspective arranges technology on a spectrum of independence to prepare readers to consider lethal weapon systems as an entity, for which we can assess moral agency and just-war applicability.

Chapter 4 investigates the concept of *moral agency* and attempts to determine if autonomous lethal weapon systems can be classified as moral agents. Moral agency is fundamental to the idea of justice, and therefore to the Just War Theory. It is shown that current philosophical definitions that inform the concept are insufficient to address all levels of the spectrum, leading to an inconclusive assessment of semi-autonomous and autonomous systems. Rather than arbitrarily choose an interpretation, we will reclassify the spectrum of independence as Agents and Non-Agents, allowing for further examination. Additionally, the concepts of *point of decision, point of consequence,* and *agency transference* are introduced as an aid in the further assessment of Just War Theory application.

Chapter 5 takes the stance that lethal weapon systems are not moral agents and uses previous concepts to analyze these "tools" of war to determine where moral culpability is held. Through agency transference, point of decision, and point of consequence, it becomes evident that Just War Theory not only still applies, but concerns with regards to the culpability of lethal weapon systems are far less pronounced because of human involvement.

Chapter 6 provides a summary of this paper's conclusions on how Just War Theory applies to lethal weapon systems. It reasserts the underlying argument that Just War Theory can be applied to current "autonomous" weapons. Even if lethal weapon systems could be considered moral agents, the inability to reward or punish such an entity precludes society's ability to hold such an entity morally accountable and, therefore, voids current Just War Theory considerations.

Chapter 2

It's Just...

> *The search for the truth is in one way hard and in another easy - for it is evident that no one of us can master it fully, nor miss it wholly. Each one of us adds a little to our knowledge of nature, and from all the facts assembled arises a certain grandeur.*
>
> - Aristotle

The ideas of justice in war and autonomous weapons are, in of themselves, complex issues. Each topic is comprised of several constituent elements; each subject has its own interpretations. When combined relative to each other, conclusions become problematic if the assertions of the argument are not built on a solid foundation. To discuss meaningfully Just War Theory's application to "autonomous" systems, we must first understand what is meant by key terms and how they relate to each other. Theory and tradition will be defined and differentiated, concepts of morality and responsibility examined, categories of the theory presented, and a focus on *jus in bello* principles will be made. In establishing a common ground for our discussion, we shall also identify areas of comparison and threads of thought which we shall weave into our central tapestry.

Just War Theory Defined

Just War Theory is a system of ideas, both philosophical and practical, that seek to explain the moral logic in how and why wars are fought. It is a theory which strives to bridge the conception of unassailable morality with its unattainability in reality. It acknowledges that, because humans are the miserable creatures St. Augustine of Hippo believed them to be, the human activity of war is inescapable. Yet St. Augustine and other influential theologians and philosophers like him believed that "sound moral norms…can

do much to lessen the worst ills of war."[1] In much the same way, the Irish Jesuit and Loyola University philosopher Father James Murphy explains "it's best to be realistic, and join together in international solidarity to impose order, using force if necessary. War is a terrible evil, in the sense of a destructive state of affairs. But that does not mean that any country going to war is acting immorally."[2] Just War Theory rationalizes war and permits one's involvement in the suffering and destruction by demanding measures of moral restraint in how and why one fights. Murphy explains that "the just war tradition is based on the idea that war ought to--and can--be used to establish a proper peace."[3] Furthermore, its grounding in international law ensures it provides a persistent framework for morally judging the concerns introduced by the use of military power.[4] Implicit in this application of morality, however, is the assertion that war-wagers and war-fighters can be held responsible for their conduct.

Tradition

The tradition of Just War is best thought of as the optimal consensus of religious and secular ideals; "articulated norms, customs, professional codes, legal precepts, religious and philosophical principles, and reciprocal arrangements that shape our judgments of military conduct."[5] It is revealed in the body of international rules and agreements, such as the Geneva conventions and Hague Rules of Land Warfare, that have guided acceptable actions in numerous wars and conflict. These rules aim at limiting certain kinds of warfare and set standards which jurists may refer to in prosecuting transgressors.

Punishing the losing side of a war is nothing new, for it indeed has been the case throughout history. Beginning with the Nuremberg Tribunal at the

---

[1] Gregory M. Reichberg, Endre Begby, and Henrik Syse, *The Ethics of War: Classic and Contemporary Readings* (Malden, MA: Blackwell Publishing, 2013), Kindle Location 50.
[2] Joe Humphreys, "Was the 1916 Rising Morally Justified?" *The Irish Times*, August 22, 2014, accessed April 10, 2018, https://www.irishtimes.com/culture/was-the-1916-rising-morally-justified-1.1902758.
[3] James G. Murphy, *War's Ends: Human Rights, International Order, and the Ethics of Peace* (Washington, DC: Georgetown University Press, 2014), 22.
[4] Bradley Jay Strawser, *Killing by Remote Control: The Ethics of an Unmanned Military* (New York, NY: Oxford University Press, 2013), 26.
[5] Strawser, *Killing by Remote*, 28.

close of World War II, such trials took on a new identity. Justice, not retribution, was made paramount. The Nazis were not on trial merely because they went to war and lost, but instead they were being prosecuted and punished for the crimes they had committed.[6] At Nuremberg, no state's body of law was used to prosecute the Nazis. Rather, the war crimes were defined by The Hague and Geneva conventions, which the Nazis had violated. Moreover, prosecutors found the strength of argument through "standards of justice that were widely held internationally - even if they were not codified into a body of 'international law'."[7] Since then, international bodies, like the United Nations' International Court of Justice, the International Criminal Court established by the Rome Statute, and various international tribunals stand on a body of treaties and international agreements to adjudicate offenses and mete out justice.[8] Philosopher and academic Bradley Strawser adds "the just war tradition, with its modern grounding in international law, persists as the best time-tested framework for thinking critically and morally about the justificatory burdens associated with resorting to military force."[9]

Theory

The theoretical aspect of Just War Theory is concerned with exploring the ethical justifications for war and its various forms. This is to say, determine how - and if - it is possible to wage war while being consistent with one's morality. Philosophers, theologians, jurists, political theorists, historians, and military strategists throughout history and all walks of life are confronted with four fundamental questions:

- "Can human beings ever take part in war without seriously violating moral obligations or destroying their moral character?
- When and under what conditions can war be rightly initiated?

---

[6] William Driscoll, Joseph Zompetti, and Susette Zompetti, *The International Criminal Court: Global Politics and the Quest for Justice* (New York, NY: International Debate Education Association, 2004), 11.
[7] Driscoll, *International Criminal Court*, 12.
[8] Roy S. Lee, *The International Criminal Court: The Making of the Rome Statute: Issues, Negotiations and Results* (The Hague: Kluwer Law International, 2002), ix.
[9] Strawser, *Killing by Remote*, 26.

- How can war be fought so that the most basic moral standards are not violated?
- What should be done to ensure a lasting peace once the hostilities are over?"[10]

In seeking the truth, as all worthwhile theories endeavor to do, Just War Theory has grown over time. It finds origins and inspiration from non-Western or non-Judeo-Christian bases such as Middle Eastern and Asian cultures, Islam, Taoism, and Buddhism.[11] Regardless of the source, Just War Theory pursues the truth about the ethical good hidden in the evils of war. The theory of Just War is not a doctrine, but a reaffirmation that good and evil are part of the more significant dialogue between what must be done and what one may do. The theory embodies humankind's aspiration to remaining morally unbroken even after the tribulations of war. Conversely, the practical aspect, or "tradition" of just war, is concerned with the manifestations of Just War Theory throughout history.

Difference

Framed differently, theory is the description of truth - or at least our best and most accurate understanding of it, while tradition is the tangible manifestation of the theory. This relationship between theory and tradition is like a newborn baby's ability to walk. In theory, so long as the child remains healthy, one may expect that she will crawl, walk, then eventually run. The fact that we have not currently observed this child doing any of these activities does not change the validity of the expectation. Years later, her consistent first-place finishes in track and field do not mean her aspiration for an Olympic Gold medal will come true. This distinction is made to differentiate and underscore our focus on the theory of Just War. Tradition aids our reasoned analysis, but we seek the truth about Just War Theory's relationship to advanced

[10] Gregory M. Reichberg, Endre Begby, and Henrik Syse, *The Ethics of War: Classic and Contemporary Readings* (Malden, MA: Blackwell Publishing, 2013), Kindle Locations 60-62.
[11] Reichberg, *Ethics of War*, Kindle Locations 64-67.

technologies, not policy. The truth may, however, naturally manifest itself in tradition.

Ultimately, it is the ethical examination of warfare that informs the creation of these covenants and checks them for their philosophical coherence and relevance to emerging conditions.[12] For example, sovereignty has long since been held as the ultimate right of a state, yet the concept of a Responsibility to Protect and the rise of transnational terrorism has forced the notion of sovereignty to accommodate outside influences due to moral justification. The information age, ubiquity of the internet, and the astronomical speed of technological revolutions gives rise to both efficient ways to bring people across the world together, and unprecedented methods to rip them apart. Humanity must again question whether its traditional assumptions governing "just" conflicts are sufficient.

Just War Theory's practical morality explains and guides the conduct of war. The theory drives men and women to take responsibility for their actions within their hearts and souls; the tradition makes them accountable in the eyes of humanity. The tradition of Just War gives one the basis of moral analysis. However, Distinguished Professor of Religion and professor of political science at Rutgers University James Johnson warns that to do so, "requires active moral judgment within a historical context that includes not only the contemporary world but the significantly remembered past."[13] No matter if one believes in the existence of a set of universal values or argues that morality is in the eyes of the beholder, Johnson reminds us that Just War Theory's representation, in reality, has been, and continues to be, subject to contextual factors.

Morality

The concepts that support and drive Just War Theory are ancient. The oldest written discussions describing justice in war can be seen in the Indian Hindu epic, the Mahabharata. Hindus regard this poem as both *dharma* (Hindu moral law) and *itihasa* (history). In it, five rulers deliberate whether the

[12] Alexander Moseley, "Just War Theory," Internet Encyclopedia of Philosophy, accessed March 11, 2018, http://www.iep.utm.edu/justwar/.
[13] James Turner. Johnson, *Can Modern War Be Just?* (New Haven: Yale, 1986), 12.

suffering caused by war could ever be justified. In their reflections, they identify essential elements of Just War Theory that are reaffirmed by later Just War theorists. Topics such as the proportionality of an attack, what weapons are permissible, intent, and the treatment of prisoners emerge.[14] Similarly, Historian Rory Cox's examination of pre-Christian just-war thought reveals that "ancient Egyptian warfare operated within a well-established system of ethics; considerations of justice were fundamental to the Egyptian concept of righteous war." There can be no mistaking that ancient Egyptian war was brutal and ugly, but as Cox explains, the effectively non-existent restraints in their conduct of war were not due to lack of thought but were a "direct consequence of the development of a prepotent *jus ad bellum* tradition."[15] The ancient Egyptians went to war believing they were doing so justly. As such, they would use whatever methods were needed to succeed. Likewise, just-war thoughts and principles are also found in Islamic teachings of Jihad, Greco-Roman doctrine, and Jewish tradition.[16]

The western understanding of Just War traditionally looks toward Catholic teachings and early theory to underwrite guiding principles. The teachings of Augustine of Hippo and Thomas Aquinas clarified Christendom's relationship to war and legitimized the idea that wars could both be waged for *good* reasons and by *good* people. In *Contra Faustum Manichaeum*, Augustine argues that Christians, as part of a state apparatus, need not be ashamed of protecting peace and punishing wickedness when forced to do so by a government.[17] Furthermore, he asserts that peacefulness in the face of a grave wrong that could be stopped only by violence would be a sin. Defense of one's self or others could be a necessity, especially when authorized by a legitimate

---

[14] Paul F. Robinson, *Just War in Comparative Perspective* (Aldershot: Ashgate, 2003).
[15] Rory Cox, "Expanding the History of the Just War: The Ethics of War in Ancient Egypt," *International Studies Quarterly* 61, no. 2 (2017): 371-384, doi:10.1093/isq/sqx009.
[16] "Jihād as Defense: Just-war Theory in the Quran and Sunnah," Yaqeen Institute for Islamic Research, accessed March 20, 2018, https://yaqeeninstitute.org/en/justin-parrott/jihad-as-defense-just-war-theory-in-the-quran-and-sunnah/.; Michael Walzer, "The Ethics of Warfare in the Jewish Tradition," *Philosophia* 40, no. 4 (2012): 633, doi:10.1007/s11406-012-9390-5.
[17] Boniface Ramsey, Roland J. Teske, and Augustine, Saint, Bishop of Hippo, 354-430, *Contra Faustum Manichaeum* (Hyde Park, NY: New City Press, 2007), 69-76.

authority. "They who have waged war in obedience to the divine command," he emphasizes, "or in conformity with His laws, have represented in their persons the public justice or the wisdom of government, and in this capacity have put to death wicked men; such persons have by no means violated the commandment, *Thou shalt not kill*."[18]

Nine hundred years later, Thomas Aquinas applied his philosophical principles of reason to build upon Augustine's allowance that Christians could be compelled to violent action in certain situations. He contends that for a war to be just, it must be waged by the proper authority, be conducted for an honest purpose, and be carried out for the right reasons. These principles of Proper Authority, Just Cause, and Right Intention have become foundational to the contemporary beliefs of Just War Theory espoused by theorists, theologians, military strategists, and political leaders of today.[19]

The theory of Just War seeks the truth of what is morally permissible, while the tradition of Just War describes humanity's attempt to infuse ethical standards into humankind's most regrettable activity and hold offenders accountable. Elements of justice are found across time and throughout the globe, but their quality and effect are primarily dependent on the beholder's sense of morality. Many believe there is no such thing as universal morality or values. Peoples of different nations and creeds hold different beliefs about the qualities and virtues they find important. In fact, even within these groups, beliefs differ in priority and character; does one speak the truth despite the pain it causes, or tell a lie and spare the heartache? One's morality is definable, and yet unreasonable in every circumstance. Morality exists in a realm Aquinas calls Eternal Laws, for faith is the key to unlocking real understanding. Even if, in truth, there is a universal set of moral values for which all are responsible, in practice there is not. This absence does not prevent the United States and her allies from denouncing or acting against states that violate fundamental human rights.

---

[18] Saint Augustine, Bishop of Hippo, *The City of God*, trans. Marcus Dods, vol. 1 (T. & T. Clark, 1888), 22.
[19] Reichberg, *Ethics of War*, Kindle Locations 2747-2748.

Aquinas asserted that *reason* was one of God's greatest gifts to humanity and that the world could be meaningfully explored through reason and not simply by faith. Through reason, humankind could acquire a breadth and depth of knowledge that was not predicated on being a Christian. In fact, this was very much the case during the Islamic Golden Age from the 8th to 14th century, when the Islamic Caliphates in Morocco, Spain, and Egypt flourished by being open to knowledge from all over the world and generated advances in scientific ideas and philosophy. This possibility stemmed from Aquinas' proposition that the world followed Secular/Natural Laws as well as Religious/Eternal Laws. Unlike God in His omnipotence, who worked through whatever manner He desired, man was limited. Aquinas believed that there would be instances, such as prophetic revelations and angelic apparitions, where God worked outside of human reason. In these instances, faith was required to understand. However, Aquinas was also adamant that humans had access to vast knowledge by applying reason within the realm of Natural Law. His application of reason, and not merely religious belief, to the ideas of Just War Theory, shows that one need not hold the same moral code as all others if, through reason, one may draw the same conclusions about proper conduct. Just War Theory is still relevant even with a perception of subjective views of morality. Justice, however, is objective because it attempts to return balance by enforcing the chosen moral code of a society; this can be the Constitution of the United States, the Sharia Law, or the community rules of a small village. Within those specific societies, justice is meted out.

Explained in a different way, the concept of morality is like the concept of time. There can be no dispute that time exists in the world. Even elements of it - the measurements of seconds, hours, days, and so on - are accepted and shared everywhere. Yet depending on one's background, situation, location, and what rules govern that place, one's interpretation of time can take on a significantly different character. The vacationing man watching the Hawaiian sunrise feels like he has got nothing but time, while simultaneously the Wall Street executive feels like she barely has a moment for lunch. The student waiting for class to end may feel like the passing of 10 minutes took hours, while the dying man wonders how 80 years have gone by so quickly. Time as a

concept, like morality, brings order to our lives and helps to guide our actions each day. We must recognize, however, that there is a difference between the rules of time and how they are applied. While each minute is 60 seconds and each hour is 60 minutes wherever one goes, the clock-time of one place may differ from another. The fact that it is 10 o'clock in the morning in Las Vegas does not mean it must be 10 o'clock in Washington DC. In fact, when it is the top of the hour in London and Beijing, in places like Newfoundland, India, Afghanistan, and parts of Australia and New Zealand it is not. The application of daylight savings time in much, but not all, of the world, further illustrates this point. Despite these differences, time itself has not changed. What changes is how each area applies it. This relationship is akin to that of morality and justice.

If morality guides one's life, justice measures how well. To be "just," then, is not simply to abide in one's values - for that is to be moral - but to apply one's moral code appropriately. Through this lens, one can see that being "just" has nothing to do with what other people think, and everything to do with what we ourselves believe. At its most essential level, moral justice is a personal affair that concerns only the judgment of the individual and, if they so believe, their God. A group's sense of justice - be it a family, tribe, nation, or alliance - is a balancing of the beliefs of constituent elements. This new sense of justice is used to create norms and laws by which everyone is judged. Even though this sense of judgment is derived from the many beliefs of many people, it has become one code owned by one unified body - the catechism of the Catholic Church, the laws of the land, and the charter of the United Nations, for example. It matters not whether an offender is a member of the group or an outsider; the Catholic, American, or organizational label applies their own code to judge the action in question. Once again one finds that being "just" has nothing to do with what other people (outsiders) think, and everything to do with what we ourselves (members) believe.

Just War Theory is relevant today and will likely continue to be important well into the future. The apparent lack of universal morality is inconsequential because Just War Theory is about reconciling the methods and purpose of war with one's values, not those of someone else. For if one has

chosen to believe in certain morals, then one must also believe one is correct in doing so. Therefore, the desire to act justly is for one's benefit and for the benefit of those who are like-minded. When one accuses another of being unjust, the implication is that they have not acted according to one's chosen moral code. This moral ascendency becomes the basis of passing judgment upon the unjust - not because they acted incongruent with their morality, but incongruent with ours.

If morality is not absolute, what grounds does Just War Theory have to stipulate what can or cannot be done? Prominent political theorist and moral philosopher Michael Walzer describes morality as a "language of justification." It is not simply a collection of normative terms directing action but also one of descriptive terms that form a lexicon from which discourse on the subject may be constructively had. He adds "the moral reality of war is not fixed by the actual activities of soldiers but the opinions of mankind."[20] The work of philosophers and lawyers and the beliefs of the public are not mere reactions to the interpretation of morality in war as isolated to battlefield events. Rather, our traditions - codified or not - continue to inform and shape the prevailing understanding of Just War. The language is a reflection of the "moral world and gives us access to it" through a vocabulary "sufficiently common and stable so that shared judgments are possible."[21] This helps to ensure that even if the world is not agreeing to aspects of Just War Theory word-for-word, we are all at least reading the same page.

We must remember that Augustine and Aquinas did not divine the ideals of justice. In fact, Aquinas' application of reason and Aristotle's discourse on justice and war with his teacher, Plato, represent a more secular conception of Just War Theory.[22] This is a focus that philosopher Brian Orend contends is at the heart of the theory.[23] To that end, Roman philosopher-statesmen such as Cato, Cicero, Julius Caesar, and Marcus Aurelius provided their reflections on

---

[20] Michael Walzer, *Just and Unjust Wars a Moral Argument with Historical Illustrations* (New York, NY: Basic Books, 2015), 13-14.
[21] Walzer, *Just and Unjust Wars*, 20.
[22] Plato, *The Republic* (John Wiley & Sons, 2012).
[23] Brian Orend, *The Morality of War* (Peterborough (ON): Broadview Press, 2013), 10.

the subject, such as Cicero's concepts of proper authority and public declaration, which continued to shape the concepts further.[24]

Responsibility

The United States' outlook on acceptable collateral damage may differ from that of France, Germany, Afghanistan, Pakistan, and other states, but still the desire to protect non-combatants exists. In the absence of a set of recognized universal values and morality, agreements like various United Nations, Geneva, and Hague conventions have been implemented to limit certain kinds of warfare and outlaw others. These agreements also provide a basis for states to declare an action unjust and find recourse for punishing war criminals and those who have committed crimes against humanity. Offenders are not only legally culpable in some cases but also can be found morally at fault in most. This is because, as Walzer and retired U.S. Army Lieutenant General James Dubik agree, soldiers can never be wholly transformed into an instrument of war - unthinking beings that obey orders without question.[25]

In late October 1942, Hitler issued his forces a Commando Order directing all enemy soldiers caught behind the German lines would be summarily executed without trial, even if in proper uniforms or if a surrender attempt was made. Erwin Rommel, one of Hitler's most famous generals, did not need the Nuremberg trials to tell him that such an order was a breach of prevailing laws of war. Rather than commit such a war crime or lead his men to do the same, Rommel burned the order and refused to comply. It is difficult to believe this man, whose allegiance was to Adolf Hitler and Nazi Germany, "could have escaped the moral infamy of the war he fought." Still, he is considered to have been an honorable man who, "while many of his colleagues and peers in the German army surrendered their honor by collusion with the iniquities of Nazism…he concentrated, like the professional he was, on 'the soldier's task of fighting'."[26] His reputation was not one shared by others in the Wehrmacht, but one of a national hero, brilliant commander, and clean fighter. He was on the wrong side of an unjust war but still fought well. The distinction

[24] Orend, *Morality of War*, 11.
[25] James M. Dubik, *Just War Reconsidered: Strategy, Ethics, and Theory* (Lexington, KY: University Press of Kentucky, 2016), Kindle Location 2924.
[26] Walzer, *Just and Unjust Wars*, 38-39.

Just War Theory makes between the decision to go to war and one's conduct while fighting reveals the possibility that a soldier might fight a just war badly or, like Rommel, fight a bad war well.

Individual soldiers can be afforded absolution in certain circumstances. The difficulty of the situations facing warriors, as Dubik asserts, may "mitigate responsibility, but does not erase it."[27] Similarly, neither can political and military leaders of any rank be entirely forgiven for their involvement in a morally questionable act. Soldiers and their military leaders are more than mere instruments. These men and women remain moral agents and are responsible for their action or inaction. Morality is vital to the theory, but one must not forget that the core of justice is judgment; one's own judgment in action and the judgment delivered upon those who have erred.

Categories of Just War Theory

Just War Theory does not exist to find an excuse to go to war. Rather, it is a call to arms to regulate, punish, and eventually remedy severe injustice. It demands leaders and warriors potentially be held to a higher standard than their enemies. They often accept more personal risk to excise society's cancers by trading swords for scalpels to minimize the damage to a world that must endure the lasting pain. However terrible war might be, it may not always be the worst option.[28] Augustine reminds us "we do not seek peace in order to be at war, but we go to war that we may have peace. Be peaceful, therefore, in warring, so that you may vanquish those whom you war against, and bring them to the prosperity of peace."[29] The principles of *jus ad bellum* (justice *of* war), *jus in bello* (justice *in* war), and more recently *jus post bellum* (just *resolution* of war) and *jus ante bellum* (just *preparation* for war), make up the contemporary concept of Just War Theory. Each group of principles might appear grammatically similar, but the focus of each is profoundly different.

---

[27] Dubik, *Just War Reconsidered*, 25.
[28] Charles Guthrie and Michael Quinlan, *Just War: The Just War Tradition: Ethics in Modern Warfare* (London: Bloomsbury, 2007), 11-15.
[29] Reichberg, *Ethics of War*, Kindle Locations 2890-2891.

Jus ad Bellum

Jus ad bellum, the justice of going to war, is a chief category of the Just War Theory. Augustine and Aquinas began the conversation by first explaining that good men may engage in violence if the proper authority wills it. That authority must, however, seek a just outcome for the right reasons if the endeavor is to be believed good. During the 1600-1900s, *jus ad bellum* seemingly fell out of importance as Hobbesian and Machiavellian realism pushed the supremacy of national *raison d'état*. The view that sovereign leaders were permitted to wage war when they felt there was a need was a result of the lack of an international authority in the anarchic world. Truly, as the Athenians in the Melian Dialogue remind us, "the strong do what they can, and the weak suffer what they must."[30]

This view gradually changed as international laws and institutions took on a pseudo-authority role.[31] Now, state leaders must ensure compliance with all *jus ad bellum* rules: competent authority (public declaration by proper authority per Orend), just cause, right intention, probability of success, last resort, and proportionality; for "failure to fulfill even one rule renders the resort to force unjust, and thus subject to criticism, resistance, and punishment."[32]

Jus post Bellum

Jus ad bellum examines the justice in a state's actions, and the newer principles of *jus post bellum* and *jus ante bellum* continue to further the discussion of an authority's obligations when resorting to violence. Although *jus post bellum* is being classified here as recent, this category of principles can trace its origin to Immanuel Kant in the late eighteenth century. Kant was keenly interested in the resolution of war and the justice of peace treaties, regime change in the defeated society, reconstruction efforts and responsibility, and elements that would aid in a lasting peace. Unlike many other philosophers from Aristotle on, who merely proclaimed the goal of war was

---

[30] Thucydides, Robert B. Strassler, and Richard Crawley, *The Landmark Thucydides: A Comprehensive Guide to the Peloponnesian War* (New York: Free Press, 2008), Kindle Locations 11849-11850.
[31] Murphy, *War's Ends*, 1-2.
[32] Brian Orend, *The Morality of War* (Peterborough (ON): Broadview Press, 2013), 64.

peace, Kant contemplated "what kind of peace can justly be imposed via war?" This far more difficult question, as Orend explains, was met with "fundamental silence, or else sweeping vagueness," which continues to persist as a "rather large number of just-war theorists completely ignore *jus post bellum*."[33] Orend takes up the mantle of *jus post bellum* thought, as the recent wars in Afghanistan and Iraq cause onlookers to wonder if the peace that was won could have been better fashioned. He allows that to victors go the spoils, but "to victors also go moral responsibilities to those they vanquish."[34]

Jus ante Bellum

In 2016, Garrett Brown and Alexandra Bohm introduced the concept of *jus ante bellum* into the debate about the proper use of force. Whereas the other Just War Theory category of principles examine the decision to go to war and resultant conditions, *jus ante bellum* asks the question, "how did it come to this?" Brown and Bohm frame this category in their criticism of international laws and commitments, like the United Nation's Responsibility to Protect. They assert that, while the international community has collectively agreed to intervene to prevent genocide, war crimes, ethnic cleansing, and crimes against humanity, this is often in the form of armed intervention. In contrast, "*jus ante bellum* proposes that if we have duties to kill in order to save distant strangers from violence, then we also have duties to alleviate the suffering of distant strangers from structural conditions that have a significant probability of leading to large-scale crisis and conflict."[35] In other words, if the world has a responsibility to *protect*, it also has a responsibility to *prevent*.

Jus in Bello

In contrast to the previous principles' focus on the justice of a competent authority's actions, *jus in bello* addresses the justice of the soldiers' violence. For the rank and file, *jus in bello* holds importance more significant than the justness of the war they are called to fight. Combat will test their strength, stamina, and skill to be sure, but will also confront them with terrible moral

---

[33] Orend, *Morality of War*, 20-21.
[34] Strawser, *Killing by Remote*, 33.
[35] Garrett Wallace Brown and Alexandra Bohm, "Introducing Jus Ante Bellum as a Cosmopolitan Approach to Humanitarian Intervention," *European Journal of International Relations* 22, no. 4 (2016): 902, doi:10.1177/1354066115607370.

decisions whose quality can never be mimicked while at peace. This personal battle will rage in their hearts and minds before, during, and after the killing. So, if good people must do objectively evil things, the rules of *jus in bello* exist to aid in mitigating the damage that must be borne.

A central foundation of Just War Theory is the assertion that war is a rules-based activity in which the rules apply equally to the warriors on either side of the fighting.[36] This point is vital to understanding if people fight well and justly, or if they do not. The rules of just fighting are believed to apply equally to both sides because of the distinction Just War Theory makes between war as a political instrument and war as a human struggle. Though states can be thrust into a violent contest for various reasons, often the ones left to die - even professional soldiers - fight because they have been told.

Walzer describes the soldier's perception of the enemy, illuminated through letters home and memoirs, with the recognition that "these human instruments are not comrades-in-arms in the old style…they are 'poor sods, just like me,' trapped in a war they didn't make. I find in them my moral equals. That is not to say simply that I acknowledge their humanity, for it is not the recognition of fellow men that explains the rules of war; criminals are men too. It is precisely the recognition of men who are not criminals."[37] This moral equality among soldiers reveals that "when soldiers fight freely, choosing one another as enemies and designing their own battles, their war is not a crime; when they fight without freedom, their war is not their crime."[38] Walzer and other Just War theorists believe that if soldiers fight other soldiers, they are not morally culpable because they have been thrust into the situation by their governments and are either executing their charge to the best of their ability or are poor sods who are merely trying to survive.

This is because *jus in bello* examines the way warriors execute their charge and against whom they cause injury. It clarifies needed restraints and limits how force may be used and is "the broad cultural consensus on appropriate limits to force that has developed over Western history…[which]

---

[36] Dubik, *Just War Reconsidered*, 9.
[37] Walzer, *Just and Unjust Wars*, 36.
[38] Walzer, *Just and Unjust Wars*, 37.

proceeds from two fundamental ideas: the need to protect noncombatants in wartime and the need for the means of war to be proportionate to the tasks of war."[39]

Orend categorizes the rules of *jus in bello* as external - rules applying to conduct against the enemy - and internal - rules that guide one's conduct in relation to one's own citizens. First and foremost, there must be discrimination between combatants and noncombatants. Soldiers are required to differentiate between legitimate and illegitimate targets, to afford civilians or institutions protections from direct and intentional attack. This discrimination is arguably the most crucial quality of *jus in bello* for it undergirds all others: benevolent quarantine of Prisoners of War - whose status as a combatant has been replaced with a form of noncombatant, due care for civilians, the Doctrine of Double Effect, proportionality, no means *mala in se*, no reprisals, prohibited weapons, and cautious and responsible use of emergent military technology. Each of these external rules of *jus in bello* has been established to ensure that the risk delivered upon those who do not have the moral equality of soldiers - that is, are not classified by a legitimate authority as an agent of death and destruction - is minimized fully within the context of war.

The internal rules are more of the same, but are focused on respecting domestic human rights owed to the population regardless whether they are citizens or citizens-turned-soldiers.[40] This was the intent of efforts such as the Paris Minimum Standards and the 1977 Second Amendment to the Geneva Convention. They recognize non-derogable human rights such as "life; not to be tortured; not to be enslaved; not to be taken hostage; 'minimal judicial guarantees'; non-discrimination; not to be subject to medical experiment; not to be subject to retroactive laws; recognition as a legal person; freedom of thought, conscience, and religion; a fair trial; a subsistence level of food and water; and special protections for children."[41]

Although this seeming litany of *jus in bello* restrictions appears to hamstring even honest efforts at victory, it is not so. Rules like proportionality,

---

[39] Johnson, *Can Modern War Be Just?,* 3, 19.
[40] Orend, *Morality of War*, 138-148.
[41] Orend, *Morality of War*, 139.

no reprisals, and prohibited weapons do not exist to prevent a decisive victory in favor of fairness. Rather, these obligations free moral fighters from the burden of evaluating every situation and determining how they may engage in battle justly. The boundaries set by Just War Theory permit fighters to focus on fighting with the knowledge that may do so without harm to their sense of morality. Further, these deontological constraints force war-wagers and war-fighters to remember that victory is meaningless if they lose themselves in the process.

Dubik's examination of Just War Theory adds a new element to the principle of *jus in bello*. He believes that through a mature understanding of the conduct of war, one must conclude that *jus in bello* pertains to more than "right conduct in the midst of battle, after the war has started," as is Orend's claim. Often, *jus ad bellum* is equated to the national strategic level, and rightly so, because that is where the decision for war resides. It follows, then, that because *jus in bello* pertains to the individual soldier, it lives at the tactical level of warfare. Dubik refutes this notion by asserting that *"jus in bello* must include not only the responsibilities of soldiers and their leaders in battle but also the responsibilities that senior political and military leaders have at the strategic level… Soldier and leader responsibilities on the battlefield are clear. The immediacy of combat provides this clarity. Less clear because of the distance from the battlefield, but no less important, is the senior political and military leader's responsibility to get decisions concerning war aims, strategies, policies, and campaigns as right as possible, then execute those decisions sufficiently well and adapt as the war unfolds."[42] Dubik here is making the distinction between the combat experienced by soldiers on the front line and the warfighting that senior leaders do from afar. Despite their physical distance from the battle, their decisions have a direct effect on the outcome of the war effort. Their inability to generate proper war aims, strategies, and the like, or to positively adapt after first and continuing contact with the enemy amount to a dereliction of their duty to the lives of the fighting men and women, and to the nation's trust that they can bring about a better peace.

[42] Dubik, *Just War Reconsidered*, 26.

Dubik's analysis reveals areas uncharted by this age-old theory of Just War. Moreover, it underscores the danger in Orend and Murphy's notion that *jus in bello* is relevant only after the fighting starts. Murphy shares this sentiment when explaining his belief in the primacy of *jus ad bellum.* In his evaluation, because *jus in bello* rules are not used except during combat, the principle is irrelevant until that point.[43] If this were true, then one could also claim that weapons of war, proficiency, and training are unimportant until the first shots of battle. However, this is not true. Weapon systems must be maintained, upgraded, and honed during peacetime, just as fighters must be organized, trained, and equipped before deployment. On-the-job training while in battle may be possible, but certainly not preferred. In the same way, nations do injustice to soldiers, sailors, marines, and Airmen if they have not prepared them to fight well. Ethical judgments are too important and carry significant tactical and often strategic implications, for it to be given a pittance of thought between trigger squeeze and impact. *Jus in bello* is not only relevant when the fighting starts - it is relevant now.

Conclusion

Just War Theory provides the conceptual framework to make sense of the appalling carnage of battle. Despite the view that Just War Theory is a Western philosophy born out of Western morality, a moment of examination reveals cross-culture origins and importance with strength of argument and clarity through secular reason. Any nation that holds itself a champion for good, like the United States, risks suffering self-inflicted wounds if it does not know how Just War Theory may shield its people. St. Augustine believed humans to be miserable creatures, and humanity's history of war concurs. If humans are not currently fighting, it is only because they are preparing for the next fight. When the time comes, one must know *how* to fight well, not only that one *can.* Weapon systems of various and increasing autonomy are rapidly coming online and are transforming the character of war in unfathomable ways. The fear that surrounds the idea of a machine autonomously choosing to end a human life and the cries of warning crescendo as the possibility of such weapons move

[43] Murphy, *War's Ends*, 2, 5.

closer to reality. When man and machine meet on the battlefield, *jus in bello* principles aid in assessing the justice of the encounter. For, in the perceived need to employ faster, technology quickly outpaces the ability to understand autonomous weapon systems and how to use them well.

Chapter 3

Machines, Independent and Intelligent...

We believe that if men have the talent to invent new machines that put men out of work, they have the talent to put those men back to work.

- John F. Kennedy

As machines become more and more efficient and perfect, so it will become clear that imperfection is the greatness of man.

- Ernst Fischer

"Autonomous" robots and artificial intelligence have become a hot topic as technological advances make science fiction closer to reality. Unfortunately, as the conversation becomes louder, so does the accompanying noise. We conflate and dilute terms used to describe and differentiate some of our most innovative and misunderstood technologies. Words like autonomous, automatic, automated, automation, artificial intelligence, independence, unmanned, pilotless, remote, drone, robot, system, remotely piloted aircraft, and unmanned aerial vehicle litter news reports, official statements, regulations, and publications when referring to weapon systems like the MQ-1B Predator or MQ-9 Reaper. In fact, Defense Secretary James Mattis shared his distaste for the term Unmanned Aerial Vehicle as "one of the most misnamed weapons in our system…It may not have a person in the cockpit, but there's someone flying it…It's not unmanned."[1] Independently, each word or phrase that we use interchangeably without thought connotes a different level of function and invokes certain biases. As complex as advanced technologies already are, our undisciplined speech creates confusion about how they can and should be used. Words matter. For our purpose of exploring Just War Theory's applicability to autonomous systems, it is critical to be clear and

[1] Richard Sisk, "Mattis' Pet Peeve: Calling Drones 'Unmanned Aerial Vehicles'," Military.com, February 21, 2018, accessed March 20, 2018, https://www.military.com/defensetech/2018/02/21/mattis-pet-peeve-calling-drones-unmanned-aerial-vehicles.html.

specific about what is being referred to as autonomous or artificially intelligent and what is actually meant by the terms. Before any meaningful discussion can be had, the misunderstandings that have warped basic terms and concepts must be stripped away to leave a lexicon that is useful and specific.

Autonomy

On the surface, "autonomy" seems to be a simple concept. Merriam-Webster defines *autonomy* as "self-directing freedom and especially moral independence," and *autonomous* as "capable of existing independently; responding, reacting, or developing independently of the whole."[2] Autonomy implies independence. According to former U.S. Army War College Associate Professor of Cyberspace Operations and Defense Transformation Chair Jeffery Caton, however, "there is no universally accepted definition for autonomy for applications involving human-machine systems." Instead, there is a general consensus that autonomy should not be thought of as a "discrete property of an object or system, but rather as a relationship between a system and its operator."[3]

The U.S. Army Robotic and Autonomous Systems Strategy defines autonomy as "the level of independence that humans grant a system to execute a given task in a stated environment."[4] The *Department of Defense (DoD) Directive 3000.09, Autonomy in Weapon Systems*, not only establishes policy and responsibilities pertaining to the development and use of these systems but similarly describes autonomous weapon systems in a manner consistent with Caton's assessment. The directive defines and differentiates these systems in a spectrum relating to the level of control afforded to a human operator: semi-autonomous, human-supervised, and autonomous. Graphically depicted below are definitions and their relationship to the human. The *Unmanned Systems*

---

[2] "Autonomy," Merriam-Webster, accessed March 20, 2018, https://www.merriam-webster.com/dictionary/autonomy.; "Autonomous," Merriam-Webster, accessed March 20, 2018, https://www.merriam-webster.com/dictionary/autonomous.
[3] Jeffrey L. Caton, *Autonomous Weapon Systems: A Brief Survey of Developmental, Operational, Legal, and Ethical Issues* (Carlisle, PA: Strategic Studies Institute and U.S. Army War College Press, 2015), 1-2.
[4] U.S. Army Training and Doctrine Command and Army Capabilities Integration Center, *The U.S. Army Robotic and Autonomous Systems Strategy* (Fort Eustis, VA: U.S. Army Training and Doctrine Command, 2017), 3.

Integrated Roadmap FY2013-2038 description of "remotely piloted" is also included to fully describe the spectrum of autonomy from direct control to autonomous. Arrows do not imply progression, but rather the ability of many systems to move from one mode to another.[5]

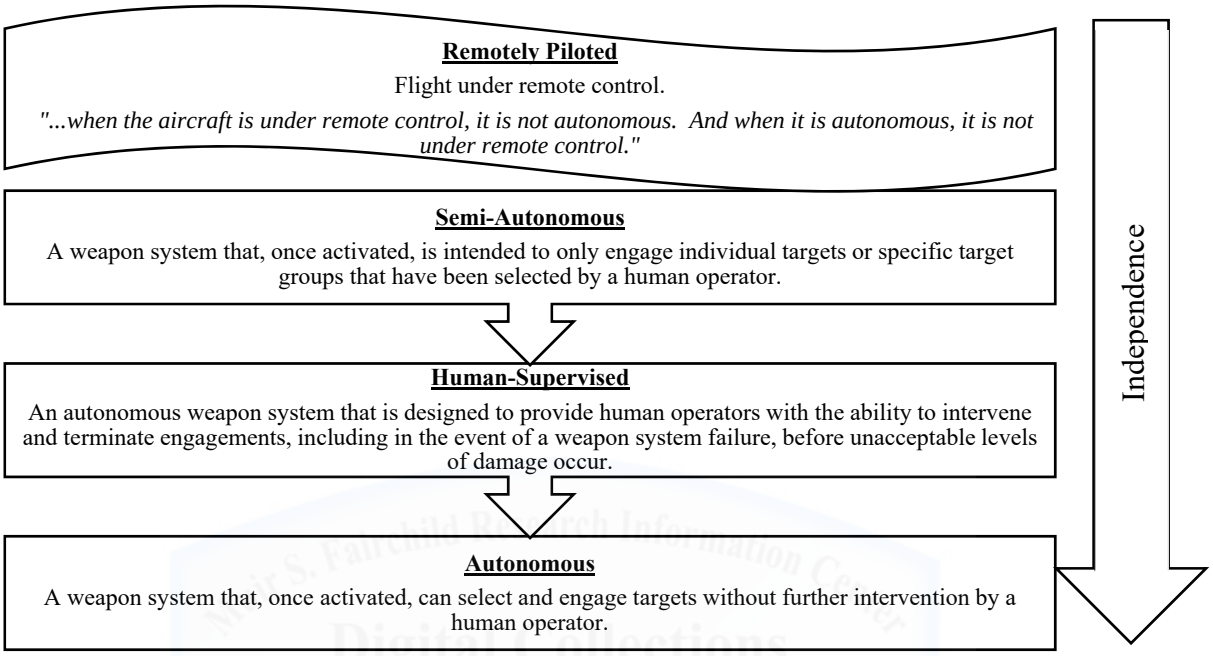

Figure 1: Spectrum of Autonomous Weapons Systems
Source: Adapted from Unmanned Systems Integrated Roadmap FY2013-2038 & DoDD 3000.09

Perhaps the clearest definition for autonomy in technology comes from the National Science and Technology Council; "the ability of a system to operate and adapt to changing circumstances with reduced or without human control."[6] This definition adds the element of adaptation to what qualifies autonomy. Still, there is a focus on the human interaction with the object. Caton's simplification of these levels of autonomy, shared by many in the Defense Department and elsewhere, continues this trend by describing "autonomous"

[5] James A. Winnefeld, Vice Chairman of the Joint Chiefs of Staff and Frank Kendell, Under Secretary of Defense (Acquisition, Technology and Logistics), *Unmanned Systems Integrated Roadmap, FY2013-2038, vol. Reference Number 14-S-0553 (Washington, DC: Department of Defense, 2013), 15.;* United States, Department of Defense, Office of the Secretary of Defense, *DoDD 3000.09: Autonomy in Weapon Systems*, by Ashton B. Carter, 13-14.
[6] National Science and Technology Council (U.S.)., Committee on Technology., Subcommittee on Machine Learning and Artificial Intelligence, *Preparing for the Future of Artificial Intelligence*, 10.

systems as "human-out-of-the-loop" while all others remain "human-in-the-loop" systems. Explicit in the directive's guidance, however, is the foundational requirement that "autonomous and semi-autonomous weapon systems shall be designed to allow commanders and operators to exercise appropriate levels of human judgment over the use of force."[7] This condition firmly joins human control and intervention to all levels of the current understanding.

"Autonomous" systems, considered human-out-of-the-loop by the directive, must paradoxically retain the human in some fashion. This requirement is understandable because of the implications that accompany a machine-initiated killing. However, if it must be so, then we do a disservice to our citizens to imply lethal human-out-of-the-loop systems exist. Rather, it is better to think of it as a human-controlled weapon in which the point of decision (to kill or operate) is made at a time and distance not coincident with the weapon's physical location (through preprogramming and logic). Even beyond the current questions of the morality of remote kinetic operations, such as "drone strikes," the current DoD guidance reveals both a reluctance to fully trust machines with lethal action and an acknowledgment that the civilian population is not ready for what it perceives as a robot making life-and-death decisions. This propensity to anthropomorphize these advanced machines because of the "autonomous" label inappropriately focuses the discussion on the suitability of a machine to "make" choices rather than the substitution of human judgment at the point of decision and consequence with algorithmic assessment.

Elon Musk (Tesla) and Mustafa Suleyman (Alphabet) are leading a group of 116 Artificial Intelligence and robotics experts across twenty-six countries who are calling on the United Nations (UN) to ban lethal autonomous weapons. Their open letter warns that "once developed, lethal autonomous weapons will permit armed conflict to be fought at a scale greater than ever, and at timescales faster than humans can comprehend. These can be weapons of terror, weapons that despots and terrorists use against innocent populations,

[7] Carter, *DoDD 3000.09*, 2.

and weapons hacked to behave in undesirable ways."[8] They do not assume cognition in autonomous weapons, but rather assert that their use will carry currently underappreciated consequences.

Former Deputy Secretary of Defense Robert Work maintains that there will always be a man-in-the-loop when it comes to life-and-death decisions, but also acknowledges that "there's so much fear out there about killer robots."[9] While the "autonomous weapons" defined by DoD are certainly a concern to many, the killer-robot fear being expressed relates to a "Terminator" scenario in which true autonomy enables a robot to "find, track and fire on targets without human supervision."[10] The *Unmanned Systems Integrated Roadmap* concedes that the "community vernacular often uses the term 'autonomy' to incorrectly describe automated operations."[11] Dr. Peter Lee of the University of Portsmouth adds clarity to the discussion by asserting "such a thing as a fully autonomous, cognizant, self-reasoning weapon system does not exist [yet]" and so much of what is argued about lethal autonomous weapon systems "is necessarily shaped by two things: perceptions of the nearest equivalents…and the influence of science fiction and the Hollywood effect."[12] As the nuanced difference is essential to a Just War Theory discussion, the spectrum of autonomy must be revisited to sufficiently differentiate our real concern - independence.

Spectrum of Independence

The previous spectrum characterizes the level of autonomy in a given platform by the amount of human involvement required to accomplish a given task. In other words, it describes the amount of automation available to a user.

---

[8] Samuel Gibbs, "Elon Musk Leads 116 Experts Calling for Outright Ban of Killer Robots," The Guardian, August 20, 2017, accessed March 20, 2018, https://www.theguardian.com/technology/2017/aug/20/elon-musk-killer-robots-experts-outright-ban-lethal-autonomous-weapons-war.

[9] Matthew Rosenberg and John Markoff, "The Pentagon's 'Terminator Conundrum': Robots That Could Kill on Their Own," The New York Times, October 25, 2016, accessed March 20, 2018, https://www.nytimes.com/2016/10/26/us/pentagon-artificial-intelligence-terminator.html.

[10] Ian Sample, "Ban on Killer Robots Urgently Needed, Say Scientists," The Guardian, November 13, 2017, accessed March 20, 2018, https://www.theguardian.com/science/2017/nov/13/ban-on-killer-robots-urgently-needed-say-scientists.

[11] Winnefeld, *Unmanned Systems Integrated Roadmap,* 15.

[12] Caton, *Autonomous Weapon Systems, 53.*

To assess Just War Theory's applicability to autonomous systems of all levels, one must look upon them as entities with limitations instead of defining them specifically by their limits. The *jus ad bellum* and *jus in bello* distinction exists because the state and the solider are looked upon as separate entities. The soldier does not inherit the moral culpability of the state for waging a bad war but rather is held to account for his own actions. The Spectrum of Independence provides this clarity through its focus on a technological unit's level of independence: its ability to self-sufficiently accomplish a mission. Each technological unit is distinguished by its maximum level of independence regardless of its capability to perform in a regressive state; an aircraft being flown with its autopilot disengaged is still an autopilot-equipped aircraft. In the Spectrum of Independence, platforms are classified as dependent, functional, executive (semi-autonomous), adaptive (autonomous), or independent.

Dependent

Dependent technology performs tasks. It is the most basic technological unit and is characterized by an absolute dependence on user input. It is incapable of independent action, requiring a user to choose, initiate, and direct all desired tasks. These technologies are often simple tools created to perform a specific purpose (cut, explode, join, contain, deflect, calculate). They are merely a mechanism to apply the user's will, and their absolute dependence means thier user's skill limits its effectiveness. Proper inputs beget intended results while wrong inputs may yield unintended results. Examples include ailerons, knives, hammers, bowls, wedges, calculators, and rudimentary digital tools.

Functionally Independent (aka automation)

Functionally Independent technology performs functions by manipulating dependent technology and managing subordinate tasks. Through automation, these platforms facilitate one or more tasks without direct user input. The user is required to choose, initiate, and direct desired functions. Once a function has been initiated, some combination of mechanical and algorithmic effort is performed automatically, manipulating dependent technology in a manner that is intended by the user but otherwise uncommanded. Examples include vehicle anti-lock brakes that, once the driver has begun braking, ensure maximum effectiveness by regulating the tasks of

disengaging and re-engaging the brakes to prevent locking; and semi-automatic and fully automatic firearms, whose construction automatically reloads the chamber with a new bullet and either prepares the weapon to fire again with each pull of the trigger or continuously repeats the firing sequence as long as the trigger is held.

Executively Independent (aka semi-autonomous)

Executively Independent technology (referred elsewhere as "semi-autonomous") is a sophisticated system that aids users in mission execution by managing functions and tasks performed by functional and dependent technologies in lieu of direct user control. User-defined missions are input via mechanical design or preprogramming prior to mission start. After a user-defined mission is chosen and initiated, either by selecting the desired system condition or consenting to automatic actions once parameters have been met, predetermined actions are executed independently of any further human interaction. Mission accomplishment - from start to finish - is unattainable solely by executively independent systems due to their inability to adapt to changing real-world conditions. A home-alarm system is an example of executively independent technology. Before mode selection, the system is programmed to act in a certain way once requisite parameters have been met. The user must engage the desired mode to initiate the alarm system's mission of monitoring; if smoke is detected it alerts the fire department and blares an alarm; or, if a glass-breaker or door sensor trips, it alerts the police department, sounds an alarm, logs events, etc. without any additional human inputs. A modern autopilot-equipped aircraft is yet another example. When engaged, the autopilot system manages the tasks of adjusting power output and control surfaces to accomplish the desired aspect of the flying mission (holding heading, course, altitude). In compliance with DoDD 3000.09, executively independent technology is interruptible by human operators.

Adaptively Independent (aka autonomous)

Adaptively Independent technology (referred elsewhere as "autonomous") is a system focused on mission performance rather than mission execution. These highly advanced systems will drive toward accomplishing desired mission goals and objectives through preferred methods but may require deviation from

their preprogrammed actions to meet mission variance.[13] These systems are highly independent with the ability to change their "behavior in response to unanticipated events during operation."[14] Even with a high degree of independence during the mission, user input is still critical through preprogrammed operational parameters, rules of engagement, decision-making matrixes, laws, and strategies, in addition to their innate functional programming. Systems may still be influenced (as desired, but not required) by user commands and guidance via system inputs or pre-programming. The strength of such a system, however, is its capability to operate independently and achieve mission success. In compliance with *DoDD 3000.09*, adaptively independent technology remains interruptible by human operators.

IBM's supercomputer "Deep Blue" gained fame in 1997 for defeating the world champion chessmaster, Gary Kasparov. Not only did a computer defeat Kasparov, but it totally changed the world of chess. Due to the "pure algorithmic intelligence" of computers, young grandmasters are training against superior artificially intelligent opponents.[15] Additionally, in May 2017, Google's artificial intelligence program "AlphaGo" defeated the "world's best player of what might be humankind's most complicated board game."[16] Both Deep Blue and AlphaGo are prime examples of adaptively independent technology. Each program was required to respond, react, and adapt to its opponent's moves rather than rigidly perform a set of pre-determined maneuvers. Even with such skill, neither program can choose to play any other game - or even to refuse to play. They were created to accomplish a certain mission - win their respective game - and that is what they will do.

Tesla's automobiles equipped with Enhanced Autopilot are currently the closest thing to fully adaptive/autonomous vehicles in the civilian sector. When on the highway with Enhanced Autopilot engaged, they "control the vehicle's

---

[13] Winnefeld, *Unmanned Systems Integrated Roadmap,* 66-67.
[14] W D. P. Watson and D. H. Scheidt, "Autonomous Systems," *John Hopkins APL Technical Digest* 26, no. 4 (2005): 368-376.
[15] Ed Finn, *What Algorithms Want: Imagination in the Age of Computing* (Cambridge, MA: MIT Press, 2017), 137-138.
[16] Paul Mozur, "Google's AlphaGo Defeats Chinese Go Master in Win for A.I.," The New York Times, May 23, 2017, accessed March 20, 2018, https://www.nytimes.com/2017/05/23/business/google-deepmind-alphago-go-champion-defeat.html.

speed based on the traffic around it, determine whether to stay in or change lanes, move between freeways, and take exits." Tesla Chief Executive Officer Elon Musk envisions a genuinely autonomous vehicle in the very near future capable of self-driving from one location to another. Still, the car is governed by Tesla's autopilot programming that enables many of the functions required for adaptivity to make a trip successful - to include restrictions on speed, the minimum distance from an object, and other rules of engagement.[17] Robotic vacuum cleaners are another example where a user can preprogram a cleaning schedule that will begin the device's function. The path that it takes and operation of the vacuum and brushes, however, are controlled by onboard sensors and factory-programmed algorithms. In either case, user input is required to initiate the function of the technology but is not required to complete the mission given.

Independent

Truly independent technology is any extremely advanced system that is capable of making its own decisions and carrying out such choices. This ability implies that such a system is capable of *reason* and, in popular discourse, may also be sentient. While systems such as these are entirely self-sufficient, their independence is not characterized by lack of human or external input - for some external stimulation is required to give them a purpose - but in their capacity to execute actions that are different, or even contradictory, to their creator's wishes. This capability is what differentiates adaptively independent technology from independent technology. Adaptively independent platforms, though able to respond to external context and determine how to execute a mission, are still bound by the programming, rules, and logic given to them by humans. Independent systems have the capacity to distill the *intent* of a mission; giving rise to the possibility that they might determine the assignment given had a low probability of success or would be counterproductive to the overall effort and choose to execute an entirely different mission. If they did

[17] Mark Matousek, "The Most Impressive Things Tesla's Cars Can Do in Autopilot," Business Insider, January 29, 2018, accessed March 20, 2018, http://www.businessinsider.com/tesla-autopilot-functions-and-technology-2017-12#musk-says-that-in-full-self-driving-mode-teslas-will-be-at-least-twice-as-safe-as-they-would-be-with-a-human-driver-at-the-wheel-15.

comply with mission orders, they might choose to disregard the Rules of Engagement (ROE) in favor of striking a High-Value-Target determined worth the cost of innocent lives, or merely interpret the ROEs differently.

To be clear - as of this writing - no such technology exists; it is strictly theoretical. Such technology is often coincident with discussions about a technological singularity and the advent of a sentient-machine apocalypse. There are three groups of people with regard to the singularity - those who do not believe it will ever happen, those who accept the possibility exists, and others who believe it is only a matter of time. The current non-existence of such technology does not mean it is unlikely ever to be, and therefore the inclusion of this level of technological independence in our examination of Just War Theory is not futile, but fertile ground for exploration as the question of real technological independence is answered in time.

Moore's law is named for the American engineer, Gordon Moore, who in 1965 claimed that the number of transistors per silicon chip would double each year. His premonition was slightly off, as reality shows that the doubling happens at about every 18 months.[18] What is important about his "law" is not his accuracy, but that it called attention to the observable exponential growth of computing power. With Google, Facebook, and other leading technology companies already using artificial intelligence to enhance their services, and the Chinese who vowed to be the leader in Artificial Intelligence by 2030 and have committed billions of dollars to the effort, it is difficult to believe humanity will stop pushing the technological boundaries of the possible in the near or distant future. We are pushing them now, and the U.S. Defense Department is giving the industry a purposeful nudge in a particular direction, if not leading the way.

It is difficult to say, with authority, that true independent technology is probable. However, it is at least possible. If it is possible that independent systems will be created, history shows that they will be delivered among a whirlwind of promise and problems as their owners strive to rapidly field the technology despite a lack of mature understanding. If independent technology

[18] The Editors of Encyclopaedia Britannica, "Moore's Law," Encyclopædia Britannica, November 29, 2017, accessed March 20, 2018, https://www.britannica.com/technology/Moores-law.

is not inherently capable of causing human suffering or death at the onset, it would be wise to expect that such a marriage of machine and mayhem will not be long to follow. Humans fear what we do not yet understand, and often we do not try to understand until that which gives us fear is here. If technology's independence is as possible as experts like noted mathematician and computer scientist, Vernor Vinge; technologist and AI expert, Louis Rosenberg; futurist and computer scientist, Ray Kurzweil; former director of the MIT Artificial Intelligence Laboratory, Patrick Winston; and the "father of artificial intelligence," Jürgen Schmidhuber believe, its current "theoretical" status is irrelevant.[19] The field of artificial intelligence is no longer marching forward in small steps but taking leaps and bounds.

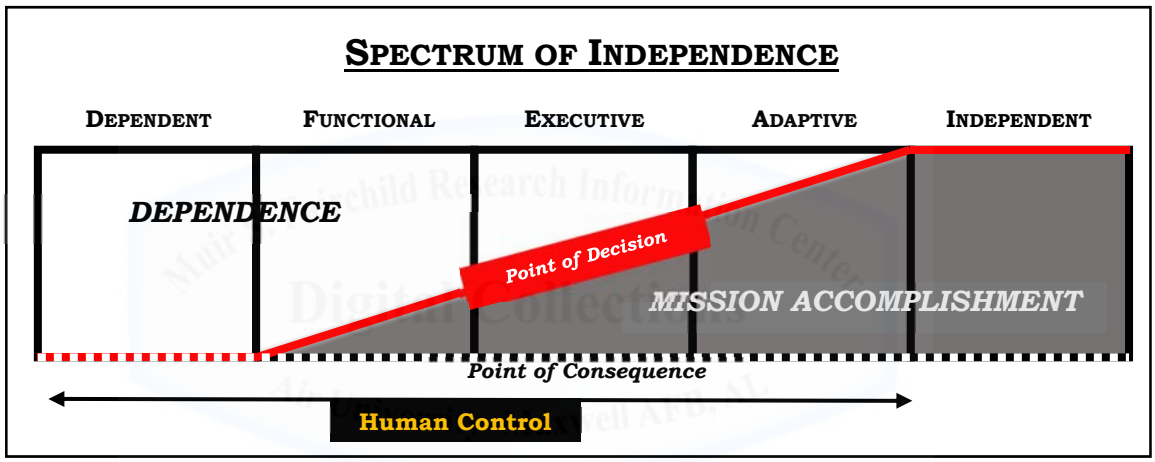

Figure 2: Spectrum of Independence
Source: Author's Original Work

Artificial Intelligence

New America Foundation strategist and senior fellow Peter Singer posits that "wrapped up in the idea of autonomy, essentially the robot's level of independence and maturity, is something even more complex: 'intelligence'."[20] Robotic intelligence is growing at an incredible pace as the field of Artificial Intelligence (AI) becomes recognized for the many benefits it may confer upon

[19] Jolene Creighton, "The "Father of Artificial Intelligence" Says Singularity Is 30 Years Away," Futurism, February 14, 2018, accessed March 20, 2018, https://futurism.com/father-artificial-intelligence-singularity-decades-away/.
[20] Peter. W. Singer, *Wired for War: The Robotics Revolution and Conflict in the Twenty-first Century* (New York: Penguin Books, 2010), 75.

humanity. Artificial Intelligence is at the very core of many executively independent and adaptively independent technologies. It will no doubt also play a large role, if not be the cause, in the creation of independent technology. Spurred on by the availability of "big data," improved machine learning algorithms, and the seemingly exponential growth of power in computers, AI can now be found in Apple's Siri voice assistant, Google Now, Amazon's Alexa and shopping recommendations, IBM's Watson computer, Tesla's self-driving cars, and a myriad of other devices that have now become mainstream.

Ironically, according to the National Science and Technology Council, there "is no single definition of AI that is universally accepted by practitioners. Some define AI loosely as a computerized system that exhibits behavior commonly thought of as requiring intelligence. Others define AI as a system capable of rationally solving complex problems or taking appropriate actions to achieve its goals in whatever real-world circumstances it encounters."[21] Additionally, the difficulty in defining AI resides in differentiating actions that can be attributed to "behavior requiring intelligence" from routine data processing. Nevertheless, the general aggregate trend in thinking points toward similar themes: a computer program that exhibits qualities of intelligence (perceived human actions/mimicry), rational problem-solving (reason), and the ability to achieve goals absent a controlled environment (adaptability through intelligence/choice). These qualities are echoed in varying degrees in our definitions of executive, adaptive, and independent technology. Ultimately, as a military analyst once told Singer, "Forget about whether the intelligence is carbon-based like humans or silicon-based like machines. Intelligence is intelligence and must be respected."[22] Therefore, more important than a specific definition of AI is understanding the capabilities of the AI that currently exists versus those which may exist in the future. These levels are commonly referred to as Narrow/Weak AI and General/Strong AI.

Current Technology

Narrow/Weak AI is the type of Artificial Intelligence most people are accustomed to, for it is the one type that currently exists. One of the primary

[21] National Science and Technology Council, *Preparing for the Future,* 6.
[22] Singer, *Wired for War*, 75.

goals of the field of artificial intelligence, as the name implies, is to eventually create an entity or program with a synthetic ability to "act appropriately (or make an appropriate choice or decision) in an uncertain environment."[23] Human intelligence can adapt to an infinite number of uncertain environments we encounter every day. If AI is to mimic this ability fully, however, programmers must also be able to code this ability to adapt. Thus far, our ability to program infinite adaptations falls short. Therefore, rather than general intelligence, AI is created for a limited, or *narrow*, set of functions "which addresses specific application areas such as playing strategic games, language translation, self-driving vehicles, and image recognition. Narrow AI underpins many commercial services such as trip-planning, shopper-recommendation systems, and ad-targeting; and is finding important applications in medical diagnosis, education, and scientific research."[24] For, as capable as narrow AI can be, it is unable to break free of its specific function; Apple's SIRI cannot drive a car, nor can a new Tesla translate German into English. Narrow AI fuels many executive and adaptive technologies, and its amazing capabilities make it seem as if our machines have a mind of their own. However valiant their effort at the imitation game, they are nonetheless not human. Narrow AI encompasses two of the four levels of AI: Reactive and Limited Memory.

LEVEL I AI: Reactive. This level of AI is the most basic level and is characterized by reactive conduct with "the ability neither to form memories nor to use past experiences to inform current decisions." Deep Blue, despite having adaptive independence, is an excellent example of reactive AI. The system's programming allows Deep Blue to recognize the pieces on a chess board and know how they move. Its algorithms give it the ability to predict possible and likely future moves for it and its opponent, enabling it to assess the most optimal move. Apart from a restriction to perform the same move three times - a very specific chess-related rule - Deep Blue disregards all actions prior to the current move. In fact, this is done on purpose, for the innovation in Deep Blue's design was not in broadening the range of possible moves, but in

[23] Singer, *Wired for War*, 75.
[24] National Science and Technology Council, *Preparing for the Future, 7.*

eliminating potential ones. The absence of "memory" prevents it from generating its own representation of the world beyond what is preprogrammed. However, in limited or explicit functions, it clearly still can perform masterly.[25]

LEVEL II: Limited Memory. The next class of AI utilizes some amount of onboard or cloud-based memory to investigate the past. This information is used to update the machine's concept of the environment and make more accurate assessments for action.[26] Self-driving cars like those by Tesla, Volvo, and Chevrolet use a mix of onboard sensors and cloud-sharing information to ensure an accurate sense of road conditions and hazards. These self-driving cars, and even less-advanced vehicles equipped only with a radar-cruise control system observes the behavior of other cars: speed and/or direction of movement. A proper reaction to another vehicle cannot be made with a snapshot of what they are doing, but rather through trend data and constant monitoring.[27] The new "memories" do not change the inherent programming of such systems, but instead constitute a library of external inputs through which more advanced function and adaptability is realized.

Theoretical Technology

As previously alluded, *general or strong AI* refers to a theoretical AI system that can adapt to an infinite number of uncertain environments. In fact, with their level of intelligence, general AI systems are expected to have the capability to problem-solve, create, and even develop personalities. Furthermore, it is this idea that gives rise to the fear that a "computer might learn so much that, at a certain point, it is not just mimicking human capabilities but has finally equaled, and even surpassed, its creators' human intelligence."[28] Independent technology will be driven, if not embody, such a strong AI. Due to the significant chasm between today's narrow-AI capability

---

[25] Arend Hintze, "Understanding the Four Types of AI, From Reactive Robots to Self-Aware Beings," The Conversation, November 13, 2016, accessed April 10, 2018, http://theconversation.com/understanding-the-four-types-of-ai-from-reactive-robots-to-self-aware-beings-67616.
[26] Hintze, Four Types of AI.
[27] T. S, "How Does a Self-Driving Car Work?" The Economist, May 12, 2015, accessed April 10, 2018, https://www.economist.com/blogs/economist-explains/2013/04/economist-explains-how-self-driving-car-works-driverless.
[28] Singer, *Wired for War*, 79.

and the expectation of general AI, both the private-sector expert community and National Science and Technology Council Committee on Technology believe that general AI will not be achieved for decades.[29] When the time comes, artificial intelligence will have a truer sense of others and potentially itself.

LEVEL III: Theory of Mind. Machines in this class of intelligence do not simply use their programmed world representation but "form representations about the world, [and] about other agents or entities in the world." In psychology, this is referred to as the "theory of mind" – "the understanding that people, creatures and objects in the world can have thoughts and emotions that affect their own behavior."[30] The ability of being able to recognize that other agents are affected by their own perceptions and motivations enables an entity with "theory of mind" conceivably to interact with humans at a social level. It could meter its behavior accordingly, for instance, to treat non-combatants, prisoners, or even other combatants humanely.

LEVEL IV: Self-Awareness. The highest level of intelligence holds the quality that is often attributed to futuristic robots of popular fiction. These machines or computer systems have sentience - self-awareness. This consciousness can be considered an extension of the "theory of mind" possessed by Type III AIs. It allows entities to be "aware of themselves, know about their internal states, and…predict feelings of others."[31] Theory of mind permits one to infer the feelings of others - "the person yelling is angry with me." Self-awareness allows the machine to acknowledge the recognition of the situation - "I know the person yelling is angry with me."

Conclusion

Independence in machines and artificial intelligence seem to go hand in hand naturally. The general concepts of each are readily reduced and understood. As seen, this general understanding is both insufficient for meaningful discourse and problematic for accuracy. The DoD's current definitions of autonomy in weapon systems may be sufficient for its needs, but they are lacking as the foundation for the application of Just War Theory and

[29] National Science and Technology Council, *Preparing for the Future,* 7.
[30] Hintze, Four Types of AI.
[31] Hintze, Four Types of AI.

the clarity of the general public's understanding. The Spectrum of Independence provides clarity by redefining the categories of autonomy through a shift of focus to the subject (the system) rather than the external influence (human involvement).

Beginning in November 2012, the Human Rights Watch (HRW) and the International Human Rights Clinic (IHRC), part of the Human Rights Program at Harvard Law School, jointly published three reports on the dangers posed by "killer robots," defined as "fully autonomous weapons that possess the ability to select and engage their targets without meaningful human control."[32] They assert that "as machines, autonomous weapons could not comprehend or respect the inherent dignity of human beings. The inability to uphold this underlying principle of human rights raises serious moral questions about the prospect of allowing a robot to take a human life."[33] Philosopher of science, technology and media Dr. Peter Asaro, echoes their worry that "in giving over the responsibility to make targeting decisions to machines, we fundamentally change the nature of the moral considerations involved in the use of violent force," and ponders "as we give over the decisions of life and death to technological systems, are we diminishing the value of human life?"[34] The concerns of Asaro, the HRW, and IHRC, disconcerting as they are, are couched in the idea that machines are making life-and-death decisions. The following chapter explores whether machines are really making decisions, or if an unseen puppeteer compels the actions of the marionette seen center stage.

[32] Caton, *Autonomous Weapon Systems,* 50.
[33] Caton, *Autonomous Weapon Systems,* 50-51.
[34] Caton, *Autonomous Weapon Systems,* 52-53.

Chapter 4

A-Moral Machine

The real question is not whether machines think but whether men do. The mystery which surrounds a thinking machine already surrounds a thinking man.

- B.F. Skinner

The question of morality in the use of lethal weapon systems (LWS) can be answered in two ways: consideration of a lethal weapon system as a tool or as a morally responsible independent actor. To determine which line of inquiry to follow, one must examine the concept of moral agency and what it means to be a moral agent. Not every entity in the world is called to distinguish right from wrong when taking action. Those that do not would not be considered praise- or blame-worthy. Those that do, are called *moral agents.* These particular agents are judged against the moral standard. Moreover, the ability to attribute responsibility for a moral or immoral action is a fundamental requirement for justice and Just War Theory. This chapter will explore what it means to be a moral agent, evaluate whether lethal weapon systems can be qualified as such, and propose concepts to aid the effort of identifying who must be served justice.

Moral Agency Defined

Unfortunately, many experts and amateurs alike have fallen prey to the imprecision of common language. The Spectrum of Independence attempts to remedy the ambiguity of the word "autonomous" and its conflation with the idea of "independence." In fact, military ethicist George Lucas Jr. maintains that "the debate thus far has been obfuscated by the confusion of machine autonomy with moral autonomy." He explains that both the Roomba vacuum cleaner and Patriot missile may be considered "autonomous" when performing their missions and adapting to obstacles and circumstance, but not in the

sense that they could change or abort their mission on moral grounds.[1] If moral autonomy (and therefore moral agency) is not commensurate with machine autonomy (system independence), one must determine what makes a moral agent and if anything on the Spectrum of Independence can be considered as such.

Agency

The first task in evaluating lethal weapon systems for moral agency is to understand the basic concept of agency and those who have it. Philosopher and lawyer Kenneth Himma describes agency, in a conceptual sense, as "the capacity to cause *actions*."[2] He continues by explaining that "actions are *doings*, but not every doing is an action."[3] These two deeds are differentiated by the mental state that one has when performing the doing. Actions are deliberate, doings may not be. For example, breathing and blinking are surely acts that one wants to continue to do, but one does not intend for the actions to happen - there is no thought required to continue the acts. The reflex to pull one's hand from a hot tea kettle or the human fight, flight, or freeze response to stress are also examples of doings that are unintended but still occur. In contrast, one may not want to get out of bed and go to work, but in doing so, the doing becomes an action because it was intentional.

Agent

Himma concludes "only beings capable of intentional states…are agents" because it is *intention* that allows one to cause actions.[4] Therefore, an entity is an agent if and only if the entity can "instantiate intentional mental states capable of directly causing a performance."[5] This represents the standard theory of agency. Additionally, agents can be natural or artificial. Natural agents are considered biological in that their existence is the result of biological

---

[1] Amitai Etzioni, "Pros and Cons of Autonomous Weapons Systems (with Oren Etzioni)," Library of Public Policy and Public Administration Happiness Is the Wrong Metric, 2018, 258, doi:10.1007/978-3-319-69623-2_16.
[2] Kenneth Einar Himma, "Artificial Agency, Consciousness, and the Criteria for Moral Agency: What Properties Must an Artificial Agent Have to Be a Moral Agent?" *Ethics and Information Technology* 11, no. 1 (2008): 21, doi:10.1007/s10676-008-9167-5.
[3] Himma, *Artificial Agency*, 19-20.
[4] Himma, *Artificial Agency*, 20.
[5] Himma, *Artificial Agency*, 21.

reproduction. Humans are prime examples of this. Dogs, whales, dolphins, and other intelligent animals are also natural agents because of their status as biological life and, arguably, their ability to do more than react to instinct. In contrast, artificial agents are "manufactured by intentional agents out of pre-existing materials external to the manufacturers."[6] These agents were not born, they were made. So long as it had the capability for intentional states, a highly advanced computer would be considered an artificial agent. KITT, the AI-equipped modified 1982 Pontiac Trans Am, from the popular 80s television series, Knight Rider, is another example.

Agency in the Spectrum

When these traits are applied across the spectrum, one finds that all types of lethal weapon systems can be classified as artificial entities capable of executing acts in the world. However, regarding "intention," evaluation becomes problematic. The standard theory, as discussed, requires one to determine if an entity is capable of an intentional mental state. The focus is clearly the "intentional" aspect, yet we find ourselves asking if an AI is capable of *any* mental state at all. The Stanford Encyclopedia of Philosophy explains that "if one takes an instrumentalist stance, there is no obvious obstacle to the attribution of mental states and intentional agency to artificial systems. According to realist positions, however, it is far from obvious…whether or not artificial systems have internal states that ground the ascription of representational mental states."[7]

Due to the current disagreement of experts regarding possible artificial mental states, it is premature to take a position on the matter. Rather, because of this evaluation, the spectrum will be separated into two groups to permit further appraisal. The first are the non-agent systems (LWS-N) of Dependent, Functional, Executive, and Adaptive platforms for which we presume no ability of intention. These levels of technology, characterized by "bullets and bombs, satellites, swords, and ships, [like] unmanned military vehicles are inanimate

---

[6] Himma, *Artificial Agency*, 21.
[7] Markus Schlosser, "Agency," Stanford Encyclopedia of Philosophy, August 10, 2015, accessed April 10, 2018, https://plato.stanford.edu/entries/agency/.

objects. They are tools. Moral responsibility applies to moral agents and not tools."[8]

The second group are the Artificial Agents (LWS-A) that include executive, adaptive, and independent systems with the presumption that it is possible to ascribe intention. The inclusion of executive and adaptive lethal weapon systems in this group is done to permit further examination of current technologies that the HRW-IHRC and others have already attributed the ability to "decide" life-or-death matters. An independent lethal weapon system, as described in the spectrum, has the ability to make decisions even if contrary to human inputs. This ability implies that an independent lethal weapon system must assign intention to its actions and therefore is capable of an intentional mental state or something equivalent. Naturally, the LWS-N group without the basic classification of agent cannot assume the mantle of moral agent. However, because another group of lethal weapon systems can be viewed as agents, the examination may continue to determine the possibility of moral agency.

Artificial and Natural Ties

Himma argues "the distinction between natural and artificial agents is not mutually exclusive."[9] As long as the definition is met, he believes, an agent that is both natural and artificial is possible. A clone might fit this requirement. He says "if we could manufacture living DNA out of preexisting non-genetic materials, then the resulting organism would be both artificial and biologically alive."[10]

In fact, synthetic DNA has been created. In 2012, National Geographic News reported the creation of new synthetic compounds called XNA. According to John Chaput of the Biodesign Institute at Arizona State University, these compounds, like DNA, can store and copy genetic information and can also be made to evolve in the lab.[11] In January 2017, scientists in the US modified

---

[8] Bradley Jay Strawser, *Killing by Remote Control: The Ethics of an Unmanned Military* (New York, NY: Oxford University Press, 2013), 34-35.
[9] Himma, *Artificial Agency*, 21.
[10] Himma, *Artificial Agency*, 21.
[11] Christine Dell'Amore, "Synthetic DNA Created, Evolves on Its Own," National Geographic, April 20, 2012, accessed April 10, 2018,

common E-coli microbes in a way that, they say, will ultimately allow them to program how the organisms operate and behave. The cells are considered a "stable form of semi-synthetic life" and "lay the foundation for achieving the central goal of synthetic biology: the creation of new life forms and functions."[12] An actual synthetic human clone does not yet exist, and the prospect itself carries significant moral questions of its own, but the building blocks of such a natural-artificial being do. In theory, if a synthetic human baby clone could be made, it would qualify as both a natural and artificial agent.

Moral Agency

Walzer reminds us "the assignment of responsibility is the critical test of the argument for justice."[13] This is judgment - by moral agents of moral agents. The intentional actions are evaluated by the moral standards of the collective, resulting in either praise or blame. In philosophy, the idea of being accountable for one's behavior is central to the standard view of moral agency. To refer to some act as moral or immoral is to imply that the actor and action both are governed by moral standards; holding moral duties and obligations. Therefore, having the condition of being beholden to moral standards, and therefore morally responsibility, is to have moral agency.[14]

Moral Agent

The primary focus is the question of moral responsibility - the blameworthiness of individuals - not their legal guilt or innocence.[15] Warfighters and war-wagers make many choices that can be qualified by thier morality. In making such choices and committing resultant acts, Walzer says that "it must be possible to single them out for praise and blame."[16] This moral

---

https://news.nationalgeographic.com/news/2012/04/120419-xna-synthetic-dna-evolution-genetics-life-science/.

[12] Ian Sample, "Organisms Created with Synthetic DNA Pave Way for Entirely New Life Forms," The Guardian, January 24, 2017, accessed April 10, 2018, https://www.theguardian.com/science/2017/jan/23/organisms-created-with-synthetic-dna-pave-way-for-new-entirely-new-life-forms.

[13] Michael Walzer, *Just and Unjust Wars a Moral Argument with Historical Illustrations* (New York, NY: Basic Books, 2015), 287.

[14] Himma, *Artificial Agency*, 21.

[15] Walzer, *Just and Unjust Wars,* 288.

[16] Walzer, *Just and Unjust Wars,* 287.

praiseworthiness in an agent identifies it as moral.[17] Put simply, "a moral agent is governed by moral standards, while the behavior of something that is not a moral agent is not governed by moral standards."[18]

To illustrate this difference, take, for example, a human child. Previous discussion shows that she is a biological entity with the ability for intentional action - she is a natural agent. Her youth and inexperience, however, have not yet produced a legitimate sense of what is wrong or right. Even with her parents teaching and guiding her, it will take time for her to understand the concepts of good and evil fully. Without this understanding, she is not held morally responsible for her actions and so is not a moral agent. Years later as a young adult with a matured sense of morality, she is held to a different standard. Her obligation to her morals now identifies her as a moral agent. This change is wholly predicated on her ability to recognize the moral standard and be beholden to it. Adults with sufficient mental disability might still be precluded from moral agency because of their inability to discern right from wrong.

Potential for Moral Agency

Now, we revisit the example of a synthetic human (clone) baby. Like the natural human baby above, they are both agents when younger but are unable to act as moral agents due to their immature sense of morality. If the human baby can grow to become a moral agent, there is no apparent reason why the baby clone could not also do the same. If the synthetic being understands good from evil and is beholden to its moral code, a synthetic human adult would be classified as an agent that is natural, artificial, and moral. This reveals that being qualified as "artificial" does not preclude an entity from having moral agency. Correlating a theoretical artificial human adult with LWS-A, however, would be incorrect. Instead, this example and train of thought is used simply to illustrate that entities need not be natural to be a moral agent. It is, therefore, theoretically possible to be artificial like LWS-A and have moral

[17] Patrick Chisan Hew, "Artificial Moral Agents Are Infeasible with Foreseeable Technologies," *Ethics and Information Technology* 16, no. 3 (May 13, 2014): 198, doi:10.1007/s10676-014-9345-6.
[18] Himma, *Artificial Agency*, 21.

obligations. However, the possibility of moral agency does not equate to certainty.

Moral Agent: Morality

Moral obligations govern moral agents. For humans, morality is believed to either be innate principles revealed through a mature understanding of the world, or developed through one's experiences. Often, regardless of its origins, one's sense of morality continues to develop and shift with added experience and exposure to things like the precepts of religion and the intent behind community laws. Where may artificial agents, like the LWS-A, find morality?

The difference between morality and ethics is a thin line of perspective; to the extent that many people do not distinguish between the two but use them interchangeably. Morality is, however, the *personal* code of conduct that guides our actions and is the basis of moral judgment. While ethics refers to an *external* code of conduct imposed on an individual by an outside entity; societal rules, military code of conduct, proper business practices of lawyers, doctors, and teachers, to name a few.[19] Thus far, there is no indication that a robot carries with it an innate sense of morality at activation. If that cannot be the source of morality, ethics might well be the answer. In fact, the field of machine ethics strives to inject rules and standards of behavior into machines in hopes of "giving machines ethical principles, or a procedure for discovering a way to resolve the ethical dilemmas they might encounter, enabling them to function in an ethically responsible manner."[20] As tricky as this prospect seems at first glance, the task of infusing a machine with ethics is much more daunting in truth. Professor of Philosophy at Dartmouth College James Moor explains that there are two ways one could conceivably accomplish the task: implicitly constrain the machine by design, or explicitly restrain its actions through programming and logic.

---

[19] Cydney Grannan, "What's the Difference Between Morality and Ethics?" Encyclopædia Britannica, accessed April 10, 2018, https://www.britannica.com/story/whats-the-difference-between-morality-and-ethics.
[20] Michael Anderson and Susan Leigh Anderson, *Machine Ethics* (Cambridge University Press, 2011), 1.

To create an implicit ethical agent, one must start with the design and function of the machine. Engineers or programmers ensure ethical constraints are met by allowing the machine to produce only ethical outcomes. Moor gives an example of an automated-teller machine and web-banking software. In both cases, due care of a person's money is expected. Accuracy in the transaction, privacy, and security are all important facets that must be dealt with by these artificial tellers. However, "a line of code telling the computer to be honest won't [work]."[21] Instead, these programs and machines are carefully planned and constructed to give out money in correct amounts, prevent third-party access to private banking information, and even include protections for those who might forget their card in the machine. This method of creating ethical outcomes from a machine does not make it an ethical or moral agent, for it is simply reacting in its nature rather than being intentionally morally about its actions.

The second method Moor prescribes is to create, or attempt to create, an ethical agent that can "do" ethics the way Deep Blue plays chess: assess the situation presented and act in the most just manner. In the same way that true artificial "choice" is thus far theoretic, clearly explicit artificial moral agents are intangible, but the basis for realizing the concept is closer than one might think. Delft University of Technology philosophy and technology experts Jeroen van den Hoven and Gert-Jan Lokhorst have blended advanced deontic, epistemic, and action logic to "serve as a bridge between ethics and a machine…suggest[ing] that a formal apparatus exists that could describe ethical situations with sufficient precision to make ethical judgments by machine."[22] Machine morality, in this sense, would be a product of advanced sensors and analysis to describe the situation to the agent; user inputs in the form of ROEs; Law of Armed Conflict considerations; definitions of basic human rights; domestic and international laws and the like to serve as the evaluation criteria; and the logic programming that allows the machine to determine the optimal solution to the ethical problem. Presumably, independent lethal weapon systems would be able to substitute programmed logic for its own, and

[21] Anderson, *Machine Ethics*, 16.
[22] Anderson, *Machine Ethics*, 17.

like the child who evaluates the external rules to which he is exposed, independent lethal weapon systems would likewise develop its own moral code.

The most prominent example of a programmed moral code is found in Isaac Asimov's three robot laws, and later, the addition of a "zeroth" law of an even higher priority. These laws prescribe that:

> *0 - "A robot may not harm humanity, or, by inaction, allow humanity to come to harm."*
>
> *1 - "A robot may not injure a human being or, through inaction, allow a human being to come to harm."*
>
> *2 - "A robot must obey orders given to it by human beings except where such orders would conflict with the First Law."*
>
> *3 - "A robot must protect its own existence, as long as such protection does not conflict with the First or Second Law."*[23]

At first blush, these laws appear to meet the intent of machine morality, for it must value and protect human life above all other functions. Unfortunately, people like Rodney Brooks of the iRobot company and roboticist Daniel Wilson consistently criticize that laws such as these are impossible to translate into actual programming. Not only must a programmer comprehend the totality of human morality - no small feat in of itself - but also be talented enough to express such concepts in meaningful programming.

Furthermore, in an article he wrote for the Brookings Institution, Peter Singer gives a reminder to all that despite consistently being touted as exemplars, these laws were merely literary elements used to drive many of Asimov's stories - tales that, ironically, told of instances where these very laws were somehow twisted to lead to undesired consequences. For example, the concept of "human" in one story was warped in such a way that robots defined only a certain type of people as human; thereby permitting the genocide of "non-human" people.[24] Admittedly, it is unfair to use fictitious laws to insinuate an impending failure of machine ethics, but the warning rings clear:

---

[23] Peter W. Singer, "Isaac Asimov's Laws of Robotics Are Wrong," Brookings, July 28, 2016, accessed April 10, 2018, https://www.brookings.edu/opinions/isaac-asimovs-laws-of-robotics-are-wrong/.
[24] Singer, Laws of Robotics Are Wrong.

however complete machine morality might be, there is always a danger of an incipient omission or a tangential element becoming the catalyst for the circumvention of safeguards.

Patrick Hew of the Defense Science and Technology Organization posits that an artificial agent cannot be morally praiseworthy until humans do not entirely supply its guiding principles and rules of behavior.[25] The core of this assertion is the differentiation of moral conduct - which may be accomplished by a robot with design or programmatic constraints - and moral responsibility which is a result of the choice to do good or to do evil. A simple reaction to external stimulus bears no intention. When the only options available to a robot are dependent on its program, the algorithmic "decision" it makes masks the absence of choice. This implies that even if experts succeed in creating an implicitly or explicitly moral system, it could *act* morally responsible but fail to truly *be* morally responsible.

If what Hew deems necessary is true, it would appear that the field of machine learning could hold some possibilities. Machine learning is a field of Artificial Intelligence in which programmers design algorithms that use processes to improve its own performance in a specific task without explicitly being programmed. Simply, a machine "learns" by testing the accuracy of a given model by applying massive amounts of data to it. Through statistical calculations, it makes incremental adjustments to the model based on the "truths" that it finds and repeats the process.[26] Many evaluations are happening in parallel, testing numerous versions of the model, discarding inaccurate versions and modifying the most accurate ones. The more information the system is exposed to, the "smarter" it gets. Eventually, so many iterations have been accomplished that the success rate for the given task has risen sufficiently that its pattern recognition permits it to make predictions.[27]

---

[25] Hew, Artificial Moral Agents, 197.
[26] Danny Sullivan, "How Machine Learning Works, As Explained by Google," MarTech Today, May 03, 2017, accessed April 10, 2018, https://martechtoday.com/how-machine-learning-works-150366.
[27] Oliver Tan, "How Does A Machine Learn?" Forbes, September 08, 2017, accessed April 10, 2018, https://www.forbes.com/sites/forbestechcouncil/2017/05/02/how-does-a-machine-learn/.

Unfortunately, because this learning is being done without human intervention, it becomes nearly impossible to understand what, how, and why the "best" algorithms work or ensure they function as desired. Data scientist Cathy O'Neil warns that even though the algorithms that power AI are mathematical, they are not inherently objective. As discussed, the more data "experience" AI is exposed to, the more fine-tuned its abilities become. O'Neil explains, however, that "the people who create algorithms decide which data matters and which should be ignored." So, a programmer's mistake or bias, therefore, could change the entire outcome of what the AI has learned, dooming it to "repeat our past practices, our patterns. They automate the status quo."[28]

In March 2016, Microsoft debuted what it intended to be a friendly chat bot called Tay. The intent of the bot was to "experiment with and conduct research on conversational understanding…learn[ing] from 'her' conversations and get progressively smarter." Unfortunately for Microsoft, Tay's twitter conversations were inundated by messages from online racists and troublemakers. Instead of casual speech patterns of the typical twitter users, Tay learned to "use racial slurs, defend white-supremacist propaganda, and even outright call for genocide."[29]

The implications of such a learning failure apply to the desire of teaching machines ethics and morality. If the seemingly impossible task of programming a robot to follow a comprehensive ethical code could be accomplished, machine learning is needed to ensure the robot developed its own sense of morality. Unfortunately, as the Tay debacle demonstrates, AI may not learn the lessons programmers intend. Humans would need to guide the process along to ensure the lessons gathered from "experience" data meet the intent - contradicting the basic notion of machine learning and undermining the attempt to develop machine morality not supplied by humans. Moreover, the lack of current understanding of how AI choose the "better" models to follow means that the

---

[28] Minda Zetlin, "5 TED Talks on AI to Watch," The Enterprisers Project, October 20, 2017, accessed April 10, 2018, https://enterprisersproject.com/article/2017/10/5-ted-talks-ai-watch.

[29] Rob Price, "Microsoft Is Deleting Its AI Chatbot's Incredibly Racist Tweets," Business Insider, March 24, 2016, accessed April 10, 2018, http://www.businessinsider.com/microsoft-deletes-racist-genocidal-tweets-from-ai-chatbot-tay-2016-3.

comprehensive ethical code provided at the start could potentially be incrementally excised from the machine altogether.

Regardless, in defining what it takes to be a moral agent, it is apparent that certain levels of technological independence do not warrant such a classification. For the LWS-A that have the potential for moral agency, morality comes either from the careful programming and design constraining the machine's ability to act inappropriately - the more daunting task of making a machine an explicit ethical agent - or guiding independent lethal weapon systems in the development of its own moral and ethical codes. To constrain a machine's activity, however, is to reduce any supposed intentional actions to merely reactions; thus negating the potential for agency. Hoven and Lokhorst's work permitting humans to "teach" a machine morality, by programming how to assess a situation accurately, apply moral standards, and select the most optimal solution, is promising. However, if a machine is taught what to do and what to think but never has the freedom to choose to act differently or, by mistake, choose incorrectly, is the machine really praise or blameworthy? The standard theory of agency asserts that without intention, it is not. If this is correct, then explicitly ethical machine agents are constrained merely by a different means. Only through independent lethal weapon systems, whose actual ability of choice gives meaning to any attempt to give it blame or praise, does the concept of moral agency take hold within a machine.

Moral Agent: Pride and Punishment

For the moment, the focus will shift from the contentious debate of the feasibility of granting lethal weapon systems any semblance of morality and assume that there is *some* standard that it may be judged against to examine its potential for reward or punishment. Accountability is the crux of agency and what gives meaning to one's moral obligations and responsibilities. Without it, compliance and noncompliance with the standard have no bearing on current or future decisions. There, in essence, would be no standard. Accountability depends on two facets: identification of the transgressor and the application of an appropriate reward or punishment. For humans, one need only look to the American Declaration of Independence for an expression of the

most critical aspects of one's humanity: Life, Liberty and the pursuit of Happiness.[30] The American forefathers acknowledge, that however absolute such rights may be, they are not assured, requiring that men and women work to secure them. Through law, order, and justice, society not only facilitates the lives of the community at large but enables various benefits to the individual citizen. When people are found to be at odds with the community - by committing deviant or damaging acts - they are punished in a manner commensurate with their offense. The inalienable rights of the offenders are compromised, as designated moral agents judge the extent of the violator's transgression from the law, then prescribe the forfeit of life, freedom, or those things which give them happiness - health, money, drugs, contact - as recompense to the wronged party. Though not in every situation, laws find many of their roots in the shared morality of the society; emphasizing positive interactions between individuals, recognition of human rights, and the prevention of otherwise immoral actions damaging to the society.

Conversely, there is nothing that can be done to LWS-A that realistically can be equated to punishment or reward. Until Level IV General AI is a reality, lethal weapon systems would have no concept of "self." Without sentience, there can be no appreciation for a reward or punishment. This sense of "self" is crucial in defining an action or event as being a benefit or penalty. The forfeit of a robot's "life" by the destruction of the unit or its "freedom" by its isolation cannot be meaningful if the entity does not realize that the action is happening to it. Philosopher and University of Victoria lecturer Angus Taylor's discussion on the moral standing of animals and animal rights parallels this situation well. Taylor explains that animals are typically denied significant moral standing because they lack the capacity to reason and "one must be the sort of being that can claim what one is entitled to and can respect the rights of others." Philosopher Robert Burch likewise posits that the purpose of rights lies in the entitlement to act a certain way in one's interest, thereby making it irrelevant to

[30] Independence Hall Association, "The Declaration of Independence: Full Text," USHistory.org, accessed April 10, 2018, http://www.ushistory.org/declaration/document/.

extend rights to those who cannot or do not desire to do so.[31] LWS-A driven by Level III AI or lower may have the ability to adapt and even, by the theory of mind, recognize other beings. However, without the understanding that it itself is a being - without self-consciousness - it, like many animals, cannot advocate for its own rights. If rights for such a being are immaterial, then to what end would the removal of such rights to life, liberty, and the pursuit of happiness attain?

For the sake of argument, consider a sentient independent lethal weapon system that values what humans value. Is it possible then to reward or punish the machine? If the independent lethal weapon system concept of "self" mirrored humans - each entity is a separate being - one surely could end its life, restrict its freedom, or even remove its source of happiness. This is a typical envisioning of sentient robots as they are anthropomorphized in science fiction literature and film. In the movie, *I, Robot,* the machine protagonist gains an awareness unlike that of the other automata, demonstrated in its desire to stay alive.[32] R2-D2, C-3PO, and BB-8 from the Star Wars franchise are further representative of robots that have, and are treated as, unique entities.[33] However, these are not the "droids" we are referring to in all situations. In other conceptions, such as the Star Trek race of the Borg, artificial entities are connected as a super-organism sharing a "hive-mind" sense of "self" more akin to ants or honey bees where the being in question is the collective rather than each individual "drone" worker.[34] The lives of the individual ants or bees are inconsequential and no more than extensions of the central mind - nodes of the computer network.

In one case, it is easy to parallel rewards and punishments and their effects between humans and machines. One can argue that individual sentient units will at least value life, freedom, and happiness, and so these values can

---

[31] Angus MacDonald Taylor and Angus MacDonald Taylor, *Animals and Ethics: An Overview of the Philosophical Debate* (Peterborough, Ont.: Broadview, 2003), 58-59.
[32] *I, Robot*, dir. Alex Proyas, perf. Will Smith, Bridget Moynahan, and Bruce Greenwood (Beverly Hills, CA: 20th Century Fox, 2004), film.
[33] *Star Wars Episode VII: The Force Awakens*, dir. J.J. Abrams, perf. Daisy Ridley, John Boyega, and Oscar Isaac (United States: Lucasfilm, 2015), film.
[34] "Star Trek Borg," StarTrek.com, accessed April 10, 2018, http://www.startrek.com/database_article/borg.

be leveraged to produce the desired outcomes. In the War Against the Machines, as imagined by the Terminator movie franchise, humans and Terminators alike will be killed in battle. The human lives lost still carry meaning to those left behind, but when one of the Terminators is destroyed, neither the machine or the SKYNET Automated Defense Network, a highly advanced self-aware artificial intelligence that controls the Terminators, cares. SKYNET simply selects a new Terminator to pursue the current mission. If instead of being destroyed in battle, the Terminator unit is forced to face trial and sentenced to execution, it again would not care.

The obvious rebuttal to this thinking is that one has not actually gone after the "moral agent" if the terminator is punished. This would be like faulting a vehicle or firearm for the death of an innocent rather than the user. In the Terminator example, is it possible to hold SKYNET accountable? In the movies, SKYNET uses the internet to spread itself to other machines and locations. In current terminology, it uploaded itself into the cloud - "a vast network of remote servers around the globe which are hooked together and meant to operate as a single ecosystem. These servers are designed to either store and manage data, run applications, or deliver content or a service such as streaming videos, web mail, office productivity software, or social media."[35] The cloud is conceptually intangible, but servers that store the data are physical and can be targeted. The sheer ubiquity and physical separation of the servers gives it a resiliency against destruction that makes even a full-scale military attack hard-pressed to be effective.

If SKYNET was housed in a smaller intranet of computers, destroying every computer might be possible and would result in the destruction of the offending program. This is not the case with the cloud. Furthermore, due to the social and economic importance of the information in the cloud, it would likely be an off-limits target, even if it were possible to eliminate all the physical data centers on earth and in space. Moreover, destruction of the cloud may succeed only in driving the malicious being into "hiding," waiting to reproduce when network connection is re-established. The WannaCry virus, for instance,

[35] "What Is the Cloud?" What Is the Cloud - Definition | Microsoft Azure, accessed April 10, 2018, https://azure.microsoft.com/en-us/overview/what-is-the-cloud/.

"infected upwards of a million machines…It expanded rapidly around the globe in mid-May, hitting hospitals, businesses and government systems." Although cybersecurity experts stopped the virus, it hasn't been eradicated. In just one month in 2017, the cybersecurity firm Kryptos Logic had to deal with over 60 million infection attempts.[36] Even with the understanding of how to stop intrusions, the comparatively basic virus program persists. How much more effective could a super-intelligent program be?

Of course, this is all conjecture, for no independent lethal weapon systems actually exist. For the technologies that do, the lack of self-consciousness of understanding prevents any meaningful reward or punishment from being carried out. In *Of Mice and Men,* the fictional character of Lennie Small is portrayed as "innocent and mentally handicapped with no ability to understand abstract concepts like death." His strength and lack of intelligence make him dangerous, as evidenced by his killing of another character whom he wanted only to stop from screaming. After the act, his concern is not the death of this person, but that his friend would be mad at him for "doing a bad thing."[37] Incarcerating someone like Lennie can protect the individual from hurting himself or others any further. Regarding guilt, however, the individual cannot be found at fault for he was not a moral agent in the first place. Furthermore, pleas of "insanity" in legal proceedings, if they can be proven, are grounds for a non-guilty verdict because of the conception that the individual was not in the right state of mind to make the intentional decision to harm.[38] For existing LWS-A technology, some actions can be taken to prevent further injury, but none that can be construed as a reward or punishment. For independent lethal weapon systems, the ability to select a viable reward or punishment is there, but one's ability to carry through is not. Even with the assumption that machine morality was possible, both aspects of justice are not

---

[36] Selena Larson, "WannaCry Cyberattacks Are Still Happening. Just Ask Honda.," CNNMoney, June 21, 2017, accessed April 10, 2018, http://money.cnn.com/2017/06/21/technology/wannacry-honda-auto-shut-down/index.html.
[37] John Steinbeck, "Of Mice and Men," Chapter 5, accessed April 10, 2018, https://www.cliffsnotes.com/literature/o/of-mice-and-men/summary-and-analysis/chapter-5.
[38] Frontline, "A Crime of Insanity," PBS, accessed April 10, 2018, http://www.pbs.org/wgbh//pages/frontline/shows/crime/trial/faqs.html.

present for LWS-A. Without both, there can be no justice or moral agency for machines.

Lethal weapon systems across the spectrum cannot be considered moral agents, either because they lack the ability for *intention* and are not agents at all; lack the moral code with which intentional actions can be judged; or because society lacks any appropriate and meaningful methods of providing reward or punishment. Nonetheless, Just War Theory can be applied to lethal weapon systems through the lens that sees these platforms as only weapons. As such, one must still be able to identify the party at fault. The concepts of *point of decision, point of consequence,* and *agency transference* aid in the investigation.

Point of Decision & Consequence

As Dubik asserts, the immediacy of combat provides clarity to the moral responsibly of soldiers and leaders on the battlefield. [39] This lucidity is owed to the *point of decision* - an action's moment of intention - and the *point of consequence* - an action's moment of effect in time and space, being near-simultaneous and co-located in most instances. For example, on 21 November 2010, in support of Operation Enduring Freedom, Lance Corporal William Carpenter stood guard at a rooftop security position of Patrol Base Dakota when it was attacked. Carpenter identified an enemy hand grenade that fell within the sandbagged position and moved to shield a fellow Marine from the deadly explosion. As a result of his actions, he was severely wounded but saved the life of his fellow Marine.[40] For his conspicuous gallantry and intrepidity, Carpenter was awarded the Congressional Medal of Honor. In the fleeting moments between spotting the grenade and the explosion, he chose to protect the life of another and acted immediately. The point of decision and the point of consequence were concurrent. The confluence of these points, though common, is not always the case.

---

[39] James M. Dubik, *Just War Reconsidered: Strategy, Ethics, and Theory* (Lexington, KY: University Press of Kentucky, 2016), 26.
[40] U.S. Army Center of Military History, "Medal of Honor Recipients - Afghanistan," Medal of Honor Recipients - Afghanistan, March 16, 2016, accessed April 10, 2018, https://history.army.mil/moh/afghanistan.html.

During the Second World War, B-17 bomber aircraft were equipped with the Norden bombsight which leveraged internal computations and autopilot controls to increase the accuracy of bombing runs. Instead of manually releasing the ordinance, bombardiers now adjusted the sighting mechanism, pilots relinquished aircraft control to the autopilot, and the computer would bring the aircraft to the appropriate position to automatically deploy the weapon. The aircrew's point of decision to drop the bomb, in this case, was the moment that control of the B-17 weapon system was relinquished to the bomb-dropping technology. The point of consequence, seconds or minutes later, was the moment of ordinance release, for the crew no longer could abort the action.

The M18 Claymore anti-personnel mine, M21 Anti-Tank mine, various Improvised Explosive Devices (IED), and other such contraptions use proximity, pressure, or other mechanisms to trigger an explosion without direct input by the user beyond placement and activation. The point of decision here was the deployment - placing and arming of the system - of the explosives. Unlike the bomber crews previously discussed who could count down to the moment of consequence, these devices can remain poised for an indeterminant amount of time. The user needs no longer be in the same vicinity or be present for the moment the technology fulfills its given mission. In these cases, as with the Norden bombsight, the point of decision and the point of consequence begin to diverge both in time and space. As if to underscore this separation, the United Nations still warns that significant tracts of land worldwide continue to be plagued by mines, even decades after their initial deployment, with countries like Cambodia and Croatia averaging over 135 landmines per square mile.[41] The points of decision to employ these weapons, having long since come and gone, still continue producing unfortunate results that profoundly change people's lives.

Though these examples show the tactical application of these concepts, Dubik's assertions of *jus in bello* responsibilities owned by strategic and operational leaders provide a complement at other levels. He focuses on the

[41] "Landmines – UNODA," United Nations, accessed April 10, 2018, https://www.un.org/disarmament/convarms/landmines/.; UNICEF, "Land Mines: Hidden Killers," Land Mines: Hidden Killers, October 15, 1995, accessed April 10, 2018, https://www.unicef.org/sowc96pk/hidekill.htm.

moral responsibilities of strategic and operational leadership to correctly identify and enact the best possible objectives and lines of effort to support the war effort and prevent wasteful suffering and loss of life. At the onset of the American Civil War, President Abraham Lincoln chose national preservation as the central war aim for the North, not the issue of slavery. This decision, political scientist Andrew Polsky notes, was "the approach that plainly commanded the broadest support." By 1862, the issue of slavery stepped back to center-stage as freeing slaves became a military necessity and Lincoln explicitly changed the North's war aims with the Emancipation Proclamation in 1863. These decisions, coupled with the Militia Act of 1862 that permitted the employment of African-descended persons in military and naval service and the Conscription Act of 1863, which established the first American draft, represent senior-leader decision-points which held persistent consequences for the war effort.[42] As military theorists like Clausewitz, Jomini, and Mahan emphasize the importance of fielded forces, Lincoln's decision to focus on preserving the union maximized Northern support for a task that would lead to the repugnant affair of brothers killing brothers. Then again, when it was evident that additional manpower was needed for the war effort, Lincoln renewed the commitment to abolishing slavery and used the momentum gained to generate more power for the cause. By current estimates, "the two armies were nearly equal in strength with less than 200,000 soldiers on each side" in July 1861; however, at its peak in 1863, Union forces totaled over 600,000 - a ratio of 2:1 against the Confederates. Two years later, the disparity grew to a ratio roughly 3:1 in favor of the Union.[43]

The confluence of the point of decision and the point of consequence on the battlefield make it easier to match a moral agent to the judgment of an action, and thus simplifies the application of justice. However, even at this tactical level of warfare, it is not difficult to find examples where the point of decision and the point of consequence are not simultaneous. As the distance between these points increase, so too does the difficulty of correlating a moral

[42] Dubik, *Just War Reconsidered*, 29.
[43] "Facts - The Civil War (U.S. National Park Service)," The Civil War, accessed April 10, 2018, https://www.nps.gov/civilwar/facts.htm'.

agent with a moral or immoral decision - though it is not impossible. Dubik's assertions and Lincoln's example reveal that the relationship between these two points also exists at the operational and strategic levels of war. Furthermore, the point of consequence is not always an *instance* of effect, but the *initiation* of effects still attributable to the original decision. Indeed, Dubik maintains that the obligation of strategic and operational leaders does not end after the first shots of battle ring out. He acknowledges that plans are rarely optimized at conflict commencement and, because of the natural fog and friction of war, leaders have a duty to control continuing consequences of previous decisions by re-evaluating war aims and efforts to make adjustments to the realities being seen. If leaders are held accountable for the far-reaching consequences of the decisions they make, tactical warriors must be held to the standard in the same way.

Agency Transference

Agents must have the capacity for intentional action. Moral, ethical, or legal agents, however, also hold the responsibility for the consequences of such action. Agency transference concerns the ability of one entity to transfer agency to another entity - that is to permit another agent to act on its behalf. The concept asserts that *authority* (ability to execute the intended action or actions) may be delegated to another agent but, regarding moral agency (and those like it), *responsibility* can never be delegated. This concept is underpinned by economic "agency theory" and its derivative principal-agent framework. This theory and framework establish the role of one entity, the agent, as the representative of the other, the principal. This relationship is established for a number of reasons, but fundamentally to have the agent represent the principal in a transaction with a third party. Economists use this framework to analyze potential problems that might manifest due to the differences in goals or desires between the two; for instance, a shareholder's (principal) desire to maximize profits vs. the company executives' (agent) need to grow the firm.[44] In principal-

[44] Musa Darayseh and Abdelaziz Chazi, "Bank Specifics, Economics Environment, And Agency Theory: Determinants of Banking Performance In GCC," *The Journal of Developing Areas* 52, no. 4 (2018): 202, doi:10.1353/jda.2018.0060.

agent parlance, an agent's compliance with the principal's desires means he is *working,* while doing anything else results in *shirking.*[45] Although the framework seeks work-maximization, agency transference focuses on the principal-agent relationship itself.

The commander-soldier, parent-teacher, and employer-employee relationships are examples of this principal-agent framework. In these relationships, one is given the authority by the other to perform some action along with an expectation that the charge will be executed; the commander empowers a soldier to kill, the parents contract the teacher to instruct their children, and the employer entrusts an employee to place supply orders, for example. This arrangement works because both entities are agents with the capability to carry out the task. Toddlers could not be expected to file their parents' taxes because they do not have capability even if they have the permission, but a police officer can provide public safety because of his training and through the authority bestowed upon him by the community. Moreover, it would be absurd to consider a hammer, painter's canvas, or stove to be an agent. While they may be critical to the accomplishment of specific tasks, these tools cannot accept the agency offered by a principal.

The hierarchical structure of the military exemplifies this relationship. The authorities received from the Commander in Chief flow to Secretary of Defense, appropriate Service Secretary, Chief, and continue down the chain of command to individual warfighters. One such authority is that of securing and ensuring the readiness of American nuclear power. In 2006, electrical fuses for the nose cone assemblies used on the Minuteman strategic nuclear Intercontinental Ballistic Missile (ICBM) were inadvertently shipped to Taiwan. This mistake was not discovered until two years later, about the time a B-52 bomber aircraft armed with six nuclear-tipped cruise missiles departed North Dakota and overflew a number of states to its Louisiana destination, during

[45] Damon Coletta, "Principal-Agent Theory in Complex Operations," *Small Wars & Insurgencies* 24, no. 2 (2013): 310, doi:10.1080/09592318.2013.778016.; Dubik, *Just War Reconsidered*, 62.

which time the pilot and crew were unaware of their devastatingly deadly munitions.[46]

As a result of the nuclear-laden cross-country flight, the Air Force punished seventy Airmen. The authority to work with nuclear weapons given to these Airmen by their leadership did not mean that they would execute their duties in the manner prescribed. In *shirking* their duties, these Airmen were guilty of "an erosion of adherence to weapons-handling standards" by substituting a complex schedule of tracking and monitoring the nuclear weapons with their own "informal" system.[47] The squadron commander responsible for munitions was relieved of duty, and many others decertified.[48] Not only were Airmen across two Air Force bases disciplined, but Secretary of the Air Force Michael Wynne and Air Force Chief of Staff General Michael Mosely were both forced to resign by Secretary of Defense Gates in part due to "a pattern of poor performance" relating to the security of sensitive military components.[49] The failure of the lowest level of agents was aggregated to a higher level and so on, blazing a path of penalty and reaching the highest-ranking military and civilian within the Air Force organizational structure. Even though the agents themselves were held accountable for their actions, so too were their leaders. As seen here, a delegation of authority from principal to agent is not the same as the delegation of responsibility. Each principal and agent held responsibilities commensurate with the authorities bestowed upon them. Therefore, each was judged according to the decisions they themselves made.

---

[46] "U.S. Accidentally Sent Ballistic Missile Parts to Taiwan," The New York Times, March 25, 2008, accessed April 10, 2018, https://www.nytimes.com/2008/03/25/world/asia/25iht-taiwan.4.11413803.html.
[47] Associated Press, "Air Force Punishes 70 Airmen for Accidental Nuke-Armed B-52 Flight," Fox News, October 20, 2007, accessed April 10, 2018, http://www.foxnews.com/story/2007/10/20/air-force-punishes-70-airmen-for-accidental-nuke-armed-b-52-flight.html.
[48] Barbara Starr, "Air Force Investigates Mistaken Transport of Nuclear Warheads," CNN, September 6, 2007, accessed April 10, 2018, http://www.cnn.com/2007/US/09/05/loose.nukes/index.html.
[49] Thom Shanker, "2 Top Leaders of Air Force Pushed Out After Inquiry," The New York Times, June 6, 2008, accessed April 10, 2018, https://www.nytimes.com/2008/06/06/washington/05cnd-military.html.

During the Vietnam War, the United States Army was involved in one of the most devasting and shameful atrocities ever to be associated with the nation, the My Lai massacre. In an apparent reprisal for the deaths of their comrades, American troops brought unbelievable carnage to Vietnamese civilians, killing over five-hundred men, women, and children over the course of three hours.[50] Many soldiers refused to participate in the slaughter despite the orders of troop commander Lieutenant William Calley, but did little more than choose not to participate. Three men, however, helicopter pilot Hugh Thompson, door gunner Lawrence Colburn, and crew chief Glenn Andreotta put themselves between the rampaging American troops and the fleeing My Lai villagers. Thompson recounts that he "never had any intention or never wanted to turn [his] weapons on Americans, but [he] was forced into it."[51] Incredibly, it took thirty years before the Army recognized these men for their moral courage.[52] Not only did the non-participant soldiers and Thompson's crew *shirk* duties placed on them, but they were morally and ethically obligated to do so. Responsibility for the massacre fell heavily on Calley. For his soldiers, responsibility for what happened manifested in varying degrees; none for Thompson's crew who protected the non-combatants, some for the soldiers whose inaction permitted the atrocity to be carried out, and responsibility proportional to Calley's for those who did the deed.

Another illustration of this transference of agency involves theoretical physicists J. Robert Oppenheimer and Albert Einstein and their involvement in the Manhattan Project. Oppenheimer was an influential force on the team developing American nuclear capabilities during World War II and is regarded as the "father of the atomic bomb." He is almost as famous for his recounting of the detonation of the first atomic bomb in the Trinity test where he recalled the words of the Bhagavad Gita: "Now I am become Death, the destroyer of

---

[50] "My Lai," BBC NEWS | Asia-Pacific | Murder in the Name of War, July 20, 1998, accessed April 15, 2018, http://news.bbc.co.uk/2/hi/asia-pacific/64344.stm.
[51] Hugh Thompson, *Moral Courage in Combat: The My Lai Story: Lecture* (Annapolis, MD: U.S. Naval Academy, Center for the Study of Professional Military Ethics, 2003), 20-21.
[52] "World | Heroes of My Lai Honoured," BBC News, March 07, 1998, accessed April 10, 2018, http://news.bbc.co.uk/2/hi/special_report/1998/03/98/mylai/62924.stm.

worlds."[53] These words relay the severity of the accomplishment and Oppenheimer's recognition that he had helped equip humanity with the ability to change the world irrevocably. In truth, he dreaded success like he dreaded failure.[54] On 25 October 1945, Oppenheimer confided to President Harry S. Truman, "Mr. President, I feel I have blood on my hands." The blood on Oppenheimer's hands was in reference to the blood of future casualties of nuclear war, rather than of the Japanese.[55] Regardless, Oppenheimer's feelings reveal the sense that even if he himself would not use the invention to kill, he held responsibility in some fashion for its future use.

In contrast, Albert Einstein had no direct contribution to the development of nuclear weapons because the U.S. government would not issue him the appropriate security clearance. Nevertheless, Einstein grew to deeply regret the small role he did play as shown in 1954 when he wrote that "[he] made one great mistake in [his] life…when [he] signed a letter to FDR [Franklin Delano Roosevelt] recommending atom bombs be made."[56] Oppenheimer and Einstein felt the moral obligations weighing heavily on their shoulders. The development and advocacy for atomic weapons were accompanied by the implication that they would be used. Even if they did not know the extent of the potential damage, they became principals to the president-agent; in the realization of atomic bombs was the bestowing of authority to the president to use them, in the ultimate realization of the danger was the multiplication of their moral responsibility.

These events reveal fundamental elements of an agent's ability to transfer authority for an intended action to another agent. In doing so, a principal-agent relationship is formed. The principal, however, cannot transfer responsibility. If an agent's authority includes the ability to delegate further, as

---

[53] James A. Hijiya, "The "Gita" of J. Robert Oppenheimer," *Proceedings of the American Philosophical Society* 144, no. 2 (June 2000): 123, accessed April 10, 2018, http://www.jstor.org/stable/1515629.
[54] Hijiya, Gita, 124.
[55] Paul Ham, "As Hiroshima Smouldered, Our Atom Bomb Scientists Suffered Remorse," Newsweek, April 12, 2016, accessed April 10, 2018, http://www.newsweek.com/hiroshima-smouldered-our-atom-bomb-scientists-suffered-remorse-360125.
[56] "Albert Einstein," History.com, accessed April 10, 2018, https://www.history.com/topics/albert-einstein/videos/einstein-regret.

is sometimes the case with the military, the authority may continue to be delegated until a principal has established a limit. Nonetheless, responsibility does not transfer to the agent. However, each agent encounters his or her own obligations in the execution of the principal's authority - be it moral, legal, or ethical. The ability of agents for *shirking* reveals that delegation of authority does not equate to the execution of authority, but neither does *shirking* imply an immoral action. The soldiers at Mai Lai each had a choice to make, and clearly, not everyone chose to comply with Calley's morally bankrupt orders. Soldiers, as agents, are capable of hesitating, dissenting, or opposing an immoral action, even if they have not the moral courage to do so. As such, Dubik maintains "it is a mistake to treat [them] as if they were automations who make no judgments at all."[57]

While the agent does not inherit the responsibility of the principal, the relationship is not reciprocal. Principals are responsible for the decision to delegate authority to chosen agents, for allowing them the ability to delegate authority further, and are ultimately responsible for the conduct of the agents acting within their authority. In other words, a delegation of authority is not a relinquishment of responsibility, but rather the acceptance of additional responsibility produced by the agent. Wynne and Moseley understand this element all too well. Finally, Oppenheimer and Einstein's feelings of guilt for the future represent the unique and personal moral obligations a principal may incur, due to the gravity of the authority given and despite the views of others. Authority and responsibility are the two commodities at the center of agency transference - the entrustment of one and the generation of the other.

Conclusion

The preceding analysis compared the requirements of moral agency to current and theoretical technology as described by the Spectrum of Independence. In this examination, it is evident that popular fears of robot killers dealing death to humans are unfounded because of the inability of lethal weapon systems to make independent decisions. Rather, they are tied to the decisions of their operators as moral agents. Although artificial moral agents

[57] Dubik, *Just War Reconsidered*, 160.

are possible, all indications lead to a conclusion that current technologies cannot function in this capacity. Even after allowing assumptions of agency, the ability to follow a moral code, and the assignment of blame, no lethal weapon system can meet the requirements needed entirely. In fact, even the theoretical independent lethal weapon system fails to conform to the demands of justice - the ability to be meaningfully rewarded or punished for actions. Though independent lethal weapon systems cannot be held to a moral standard, their sentience makes them more than tools. They are beings - not unlike animals, children, or the mentally infirmed - who are not morally accountable but whose life is deserving of respect and protection.

Without the status of moral agent, robots are seemingly incompatible with Just War Theory. However, the concepts of point of decision, point of consequence, and agency transference help to organize the relationships between similar and dissimilar agents. In applying these ideas to instances where lethal weapon systems ostensibly are the cause of inadvertent deaths, one can discern where moral agency - and fault - lies.

To this end, Strawser offers a different perspective on the ethical conundrum of lethal weapons. Instead of morality, he advocates for reliability, because "like our rifles, missiles, undersea torpedoes, and jet aircraft, we demand that our military robots…won't malfunction and inadvertently destroy the operator or creator, or wantonly destroy property or innocent human life." The desire for greater independence in machines is understandable because of the efficiency and force-multiplicative effects such systems provide. Yet "they cannot, nor, in the final analysis, do we need to have them, 'behave ethically.' That is certainly asking for much more than we can currently deliver, and probably much more than we really require of them."[58] If moral agency cannot be transferred to a machine, then the point of (moral) decision stays with the human operator at some level. So long as the lethal weapon system will operate in the manner one expects, human moral agents need only face the moral implications of the decisions they make instead of also contending with those caused by the malfunctioning of the gears of war.[59]

---

[58] Strawser, *Killing by Remote*, 225.
[59] Epic Games and Microsoft Studios, Gears of War, Xbox 360 game, November 7, 2006.

Chapter 5

The Gears of War

The most powerful weapon on earth is the human soul on fire.
-Ferdinand Foch

What's in a name? That which we call a rose
By any other name would smell as sweet.
- William Shakespeare

Ethicist Robert Sparrow stresses that "any weapon or other means of war that makes it impossible to identify responsibility for the casualties it causes does not meet the requirements of *jus in bello*, and, therefore, should not be employed in war." Sparrow is referring to AI-equipped executive and adaptive lethal weapon systems with the perceived ability to make decisions on their own, leading to a difficulty in determining whether a seemingly immoral decision is due to design, program, or logic flaws.[1] What he describes, AI decision-making, is in the realm of the theoretical independent lethal weapon system. Still, his concerns are valid.

Without the capacity for moral agency, lethal weapon systems cannot be considered as making decisions and independently perpetuating violent acts. If one believes machines are not moral agents, then, as Sparrow asks - who is to blame? Far from being merely an academic question, the Human Rights Watch and International Human Rights Clinic declared that because "these robots would be designed to kill, someone should be held legally and morally accountable for unlawful killings and other harms the weapons cause."[2] Additionally, identifying the moral actor culpable for an action is only part of the issue. The question of suitability with regard to the principle of distinction

[1] Amitai Etzioni, "Pros and Cons of Autonomous Weapons Systems (with Oren Etzioni)," *Library of Public Policy and Public Administration Happiness Is the Wrong Metric*, 2018, 257, doi:10.1007/978-3-319-69623-2_16.
[2] Bonnie Lynn Docherty, *Mind the Gap: The Lack of Accountability for Killer Robots* (New York, NY: Human Rights Watch, 2015), 37.

must be considered, or else the use of lethal weapon systems would again fail to meet the requirements of *jus in bello.*

Secret (Moral) Agent - Man

This paper argues lethal weapon systems are, all cases current and theoretical, incapable of moral agency. Not only may agency transference explain the relationship between two entities with similar agency, but it can also provide insight into the relationship of dissimilar agents like that of a warfighter and a lethal weapon system. Together with the ideas of the point of decision and point of consequence, one may identify the moral agent at fault when a lethal weapon system is accused of immoral conduct.

As previously discussed, the prevailing beliefs about executive and adaptive lethal weapon systems contend that they are independently choosing targets and making life-and-death decisions about humans. In essence, the concern is that humans have abdicated their moral responsibility and transferred moral agency to a machine. As previously addressed, however, this is not possible due to the current technologies' incompatibility with moral agency. Without being moral agents, there can be no transference of moral agency to a machine; yet the transference of agency is possible.

The Predator and Reaper Remotely Piloted Aircraft saw a significant rise in use during Operation Enduring Freedom. As more and more Hellfire missile strikes were being conducted in support of counter-terrorism efforts, the controversy surrounding these flying robots rose amidst growing collateral-damage concerns. To address the controversy, internationally recognized legal expert Michael Schmitt published a detailed legal analysis in which he concludes "there is little reason to treat drones as distinct from other weapons systems with regard to the legal consequences of their employment." The consideration of these executive lethal weapon systems as indistinct from other weapons is telling with regard to the moral implications of their use as well.

Schmitt's scrutiny resulted in two additional perceptions as related to the use of RPAs: because of "their capabilities for long duration loiter and precision strike, drones may in some cases be the most legally responsible choice of force application" and that operators remotely flying the drones in the United States

are considered legitimate targets of enemy attack.[3] This last observation speaks directly to where moral agency is held. As legitimate targets, agency is followed back to RPA operators who make the decisions (not the platform) of how the weapon system is operated and have final say if and when a strike will be made.

In most cases of lethal employment by an RPA, the pilot's point of decision and consequence are separated by mere seconds. In fact, more often than not, the crew retains custody of the missile until impact, permitting an abort if required. In this situation, the RPA - despite being operated thousands of miles away - is no different from the soldier's rifle or marine's grenade. There is no transference of agency and therefore no question of where moral agency remains.

Due to their heavy use in America's Global War on Terror, RPAs have captured the attention of many, but their man-in-the-loop configuration scratches only the surface of lethal weapon systems. Peter Singer already sees a redefinition happening in what constitutes having humans "in the loop." More and more, as Drew Bennett of iRobot describes, humans will still be "in the loop," but the loop will be much wider. Futurist Ray Kurzweil similarly taunts that man might still believe he is in control, but only at certain levels.[4]

During the Second Gulf War in 2003, U.S. Patriot missile batteries mistakenly classified two allied planes as Iraqi rockets. From the moment the "rockets" first appeared there were only a few seconds to decide. Based on their trust in the system, human controllers chose to open fire and shoot down the aircraft. Singer reduces their role "in the loop" to simply a veto power. However, he continues, "even that was a power they were unwilling to use against the quicker (and what they viewed as better) judgment of a computer."[5] Does moral responsibility have the same quality in this situation? Absolutely. It is easier to feel compassion for the Patriot crew after hearing of the system malfunction, but the point of decision did not occur in the split seconds between rocket misidentification and launch. Rather, the point of decision to

---

[3] Jeffrey L. Caton, *Autonomous Weapon Systems: A Brief Survey of Developmental, Operational, Legal, and Ethical Issues* (Carlisle, PA: Strategic Studies Institute and U.S. Army War College Press, 2015), 40.
[4] Peter W. Singer, *Wired for War: The Robotics Revolution and Conflict in the Twenty-first Century* (New York: Penguin Books, 2010), 125.
[5] Singer, *Wired for War,* 125.

fire the Patriot on inbound rockets was made well before the allied planes even departed earlier that day.

Singer's assessment of the controllers' reduction to veto power stemmed from the trust the crew developed for the computer's ability. Had the allied planes been actual rockets, the crew would have performed the same way and been heroes. Had rockets been misidentified by the computer as allied aircraft, they would have failed to shoot down the threat. Although it would be unfair to fault the crew for trusting in their weapon system - warfighters cannot do their mission without having faith that their equipment will work as advertised - it remains that the point of (moral) decision was reached when the crew decided to trust the computer. The instant the "rockets" appeared on the screen signified the point of consequence. The "choice" to shoot down the rockets was an illusion. In effect, it was a mere reaction to the previous decision to trust the computer in an urgent situation. There was, of course, no transfer of moral agency from the operators to the Patriot system because a Patriot cannot be judged against the moral standard. One may argue, however, that authority to launch, in this case, was delegated to the machine. Rather than using the automatic firing mode available with the Patriot, the controllers relegated themselves to be the Patriot's human firing mechanism. No matter the level of control or authority given to the lethal weapon system, the potency of such is a direct result of human action and inaction. The choice was theirs and so too the moral obligation.

Perhaps one of the most infamous incidents involving a lethal weapon system resulted from the Aegis-equipped U.S.S. *Vincennes*, affectionately called the Robo-cruiser. The Aegis Combat System was first introduced in the 1980s to help defend Navy ships against air, surface, and submarine threats through its advanced command-and-control and weapon-control system. Aegis provided users with four modes: "Semiautomatic, in which the humans interfaced with the system to judge when and at what to shoot; Automatic Special, in which the human controllers set the priorities, such as telling the system to destroy bombers before fighter jets, but the computer then decided how to do it; Automatic, in which data went to human operators in command, but the

system worked without them; and Casualty, where the system just did what it thought best to keep the ship from being hit."[6]

While patrolling in the Persian Gulf, the Robo-cruiser's Aegis identified an inbound aircraft and displayed it with an icon implying it was an Iranian F-14 enemy fighter. Even though other data available indicated that plane was not a fighter jet, "they trusted what the computer was telling them more. Aegis was on semiautomatic mode, but not one of the eighteen sailors and officers on the command crew was willing to challenge the computer's wisdom." As in the Patriot situation, the crew authorized it to fire. Instead of two downed allied aircraft, however, a single passenger plane, Iran Air Flight 655, with 290 souls on board, were shot out of the sky.[7]

This example continues to show the trend that humans, though remaining in the kill chain, show deference to the capabilities of lethal weapon systems. Instead of humans being the center of lethal actions, lethality itself has upstaged humans. Warfighters have allowed their role to diminish from directors to participants, or worse - spectators. As compared to the Patriot crews who could merely react, Captain Will Rodgers III of the *Vincennes* made the conscious decision to authorize fires. Rodgers transferred agency to the Robo-cruiser, but it was not the malfunctioning vessel or weapon system that killed those people. Rodgers' decision was the cause. This is not to imply that this tragedy lies solely at the feet of Rodgers. That the he had the authority to fire at all without seeking permission from more senior officers in the fleet was a direct result of the command's trust in the Aegis system. In fact, Rodgers' ship was the only ship with that authority, further signaling greater trust in the computer "than any human captain's independent judgement on whether to shoot or not."[8]

All else being equal, had the Vincennes been in Automatic Special or Automatic mode, where the computer worked with limited or no human input to defend the fleet, the misidentified F-14 would still have been shot down. The moral responsibility, likewise, would not have changed. At the authority of the

---

[6] Singer, *Wired for War,* 124.
[7] Singer, *Wired for War*, 125.
[8] Singer, *Wired for War,* 124.

fleet's leadership, Rodgers' Aegis-equipped ship was still able to fire without senior-officer approval, Rodgers (principal) would have delegated the authority to the Aegis system, and the Aegis (agent) would have fired. In this counterfactual where Rodgers made no intentional decision to fire after the fighter aircraft alert, Rodgers' point of decision occurred when he gave authority to the computer by selecting an automatic mode. Though the point of consequence does not occur until Iran Flight 665 is in the air, it remains that Rodgers' decision leads to the mishap.

Without moral agency, lethal weapon systems are merely the instruments of war that Dubik asserts human soldiers could never be reduced to being. Yet even the tools of war must conform to the rules of war, or be cast aside. Like Sparrow's concern about the ability to hold the moral agent accountable, the use of lethal weapon systems must not indiscriminately harm noncombatants or be counted immoral.

Distinction/Discrimination

The status of lethal weapon systems as mere weapons of war does not preclude the onus on humanity to ensure that the instruments chosen for battle permit the protection or exclusion of those not involved with the fighting. This idea of *noncombatant immunity,* also referred to as *distinction* or *discrimination*, Johnson explains, "has been one of the strongest and most regular themes in just war tradition throughout its development."[9] Moreover, it is the basis of the assertion of the need to develop "weapons usable in ways that satisfy legitimate military functions without corollary damage to the lives, livelihoods, and property of noncombatants."[10] Other Just War theorists agree. Orend emphasizes that it is the most important *jus in bello* rule, compelling soldiers to "exert every reasonable effort to discriminate between legitimate and illegitimate targets." Anyone or anything engaged in harming would be considered a valid target - uniformed soldiers, mercenary forces, insurgents, and the like.[11] Currently, there is a debate about the suitability, accuracy, and

[9] James Turner. Johnson, *Can Modern War Be Just?* (New Haven: Yale, 1986), 27.
[10] Johnson, *Can Modern War be Just?,* 28.
[11] Brian Orend, *The Morality of War* (Peterborough (ON): Broadview Press, 2013), 112-113.

reliability of executive and adaptive lethal weapon systems' ability to differentiate between valid and invalid targets. Thus, highly regarded computer scientist Noel Sharkey "has called for a ban on 'lethal autonomous targeting' because it violates the Principle of Distinction."[12] The following examination seeks to evaluate the validity of Sharkey's claim.

Walzer explains that an "armed man trying to kill me 'alienates himself from me…and from our common humanity' and in so doing he forfeits his right to life." Regardless whether such forces are arrayed against another because of a just cause or not, the *prima facie* of this situation makes a strong case that soldiers targeting other soldiers is permitted.[13] This concept that Walzer presents is widely accepted and referred to as the moral equality of soldiers (MEOS).[14] MEOS distinguishes the combatants' right to kill in war between two groups: those they may kill and those whom they may not.

In this discussion, two questions emerge - how does an independent lethal weapon system-being conform to the MEOS, and can lethal weapon systems be restricted sufficiently to permit discrimination by the moral agents who employ them? The MEOS stipulates that combatants have given up their right to life in exchange for the right to kill. If independent lethal weapon system-beings fight, kill, and die, it follows that they should also be considered combatants and subject to the same considerations. Independent lethal weapon system-combatants then, like human soldiers, may be killed regardless of the side for which they fight.

Despite MEOS' wide acceptance, Orend and Oxford University Professor of Moral Philosophy Jeff McMahan question its validity. They argue that the "soldiers fighting for a just cause haven't given up their right not to be killed." Like policemen called to defend against criminals, their status does not strip them of the expectation to live. When they kill, they do in self-defense or in defense of those who cannot defend themselves from the threat. However, McMahan contends, "fighters in an unjust war cannot claim self-defense any more than a bank robber can claim self-defense for shooting the armed guard

[12] Etzioni, Pros and Cons, 256.
[13] Orend, *Morality of War*, 116.
[14] Orend, *Morality of War*, 113.

defending the bank he's trying to rob."[15] Though if this be the case, who is to say which side's cause is just? Were American revolutionaries just in demanding independence from England, or was King George III just in his attempt to preserve his nation? What of Lincoln's "war of Northern aggression" against the South's secession to preserve its way of life? Further discourse among philosophical experts is needed to determine the exact nature of soldiers' moral equality or inequality, but for now, it is sufficient to investigate whether Orend and McMahan's line of thinking rules out the use of lethal weapon systems.

Under this conception of the moral standing of combatants, a just independent lethal weapon system, one that fights for the side with just cause and maintains a right to life, is no different from a human combatant. Its status as a non-moral agent is irrelevant because its classification as a just combatant permits it the ability to kill an unjust enemy so long as no other rules of war are violated in the act. Further, an independent lethal weapon system killed in the line of duty produces a comparable moral obligation for the unjust soldier because he was never justified in taking any life - biological or not. If the particular independent lethal weapon system in question is not killed but just stopped - for instance a drone representation of the collective mind - the blame received would be for the destruction of property and not the taking of life; like the destruction of a tank but not of its crew.

In the same way, unjust independent lethal weapon systems fighting other combatants incur the same obligations as their unjust human counterparts. They have no "right to life" because they are among the aggressors. Therefore, blame is incurred when aggressive acts are committed, but their destruction causes none for the opposing just force. Blame when it is received, however, is held by the human moral agent that set it on its mission. In either Walzer's interpretation or Orend/McMahon's interpretation, the fact that a combatant is a human or an independent lethal weapon system does not change the applicability for this element of distinction.

---

[15] Bradley Jay Strawser, *Killing by Remote Control: The Ethics of an Unmanned Military* (New York, NY: Oxford University Press, 2013), 33.

The distinction of those who may *not* be killed poses far more complexity to lethal weapon systems. The debate revolves around the expectations that the AI directing lethal weapon systems will not only be able to differentiate between a combatant or noncombatant, but also improve the accuracy of those assessments. Despite the disagreement between Orend, Johnson, and Walzer about who is causing harm, they agree that "all those not engaged in creating harm" must not be subject to attack.[16] Discrimination of combatants and noncombatants through AI is a contentious topic. Some believe that using AI will make the task easier because computers are faster in processing large amounts of data and can perform rational unbiased analysis, while others contend that AI makes too many mistakes - even if they could get past the notion that a computer is designating a target. Militaries have long used technology to make determinations on the status of unknowns on the battlefield, be it the Identification Friend or Foe (IFF) system onboard aircraft, early-warning radars, satellite collections, signals intelligence, geospatial intelligence, measurement and signature intelligence, or electronic intelligence. Surely there are still humans in the loop who are involved in the assessment, but their propensity to favor digital analysis - especially for exquisite data - at least mirrors that of the operators on the front.

The lure of technology's promise is embodied by the DoD's Third Offset Strategy formulated in November 2014 by Defense Secretary Chuck Hagel. This strategy refocuses the DoD's innovation efforts to address the growing conventional-countermeasure capabilities of rival states.[17] Technologically, then Deputy Secretary of Defense Robert Work provided five areas of Third Offset focus: "autonomous learning systems, human-machine collaborative decision-making, assisted human operations, advanced manned-unmanned systems operations, and network-enabled autonomous weapons and high-speed projectiles."[18] The intent of such a push is to boost technological capabilities to such a degree that adversaries are incapable of producing countermeasures.

---

[16] Orend, *Morality of War*, 116.
[17] Jesse Ellman, Lisa Samp, and Gabriel Coll, *Assessing the Third Offset Strategy*, report, CSIS International Security Program, Center for Strategic & International Studies, vol. March 2017 (Washington DC: Center for Strategic & International Studies, 2017), 1.
[18] Ellman, *Assessing the Third Offset*, 3.

One avenue of investment in advanced technologies is manifested in Project Maven. In April 2017, Work launched Project Maven to "accelerate the department's integration of big data, artificial intelligence, and machine learning into DoD programs." Its intent, as explained by the director for Defense Intelligence for Warfighter Support in the Office of the Undersecretary of Defense for Intelligence Lt Gen John Shanahan, "is to turn the enormous volume of data available to DoD into actionable intelligence and insights."[19] In essence, the project aims to find a way to harness the plethora of unexploited information available to help with the accuracy of identifying military targets and discerning adversary intent. Marine Corps Colonel Drew Cukor, chief of the Algorithmic Warfare Cross-Function Team in the Intelligence, Surveillance, and Reconnaissance Operations Directorate-Warfighter Support in the Office of the Undersecretary of Defense for Intelligence, previews that the capability which Maven may enable is "computer vision -- an aspect of machine learning and deep learning -- that autonomously extracts objects of interest from moving or still imagery."[20] If successful, the ability to quickly acquire pertinent information from picture and video would significantly accelerate every stage of the F2T2EA (Find, Fix, Target, Track, Engage, Assess) dynamic-targeting process. When coupled with a system like Amazon Rekognition - an AI web service that "allows software developers to quickly and easily build applications that analyze images and recognize faces, objects, and scenes" - and adaptive lethal weapon systems, the United States would have a fearsome platform that could accurately and effectively eliminate enemies of the state while minimizing attacks on unintended targets - thus facilitating distinction.[21]

---

[19] Cheryl Pellerin, "Project Maven Industry Day Pursues Artificial Intelligence for DoD Challenges," U.S. Department of Defense, October 27, 2017, accessed April 10, 2018, https://www.defense.gov/News/Article/Article/1356172/project-maven-industry-day-pursues-artificial-intelligence-for-dod-challenges/.

[20] Cheryl Pellerin, "Project Maven to Deploy Computer Algorithms to War Zone by Years End," U.S. Department of Defense, July 21, 2017, accessed April 10, 2018, https://www.defense.gov/News/Article/Article/1254719/project-maven-to-deploy-computer-algorithms-to-war-zone-by-years-end/.

[21] "Amazon Rekognition Demo for Defense," Amazon Web Services, November 06, 2017, accessed April 10, 2018, https://aws.amazon.com/blogs/publicsector/amazon-rekognition-demo-for-defense/.

Furthermore, Professor Ronald Arkin of the Georgia Institute of Technology offers "several root causes of war crimes attributable to human motives (such as revenge, dehumanization, immaturity, frustration, and pleasure) that could be improved by the use of automated systems."[22] Not only may lethal weapon systems acquire targets with machine efficiency, but they may also negate the human conditions that cloud moral judgments during strikes. For example, the attack is a war necessity authorized by command - not one for revenge or for love of the kill.

Not everyone is thrilled about the prospects of efforts like Project Maven or the uncompromised poise of killer robots. In April 2018, over three-thousand Google employees and some senior engineers signed a letter protesting their participation in Project Maven. Despite Google's involvement in the project scoped for "non-offensive" uses, the video analysis Google AI seeks to improve would undoubtedly aid in counterinsurgency and counterterrorism operations.[23] Beyond the idealistic concerns of Silicon Valley, however, the onboard ability of lethal weapon systems to execute distinction is, at a minimum, problematic.

Cathy O'Neil's warnings about the biases inadvertently introduced into AI algorithms make the prospect of identifying those who may not be killed more complicated than it already is. This amplifies the insurgent strength of being able to hide in plain sight, often without uniforms and acting like true noncombatants until they are ready. A computer would need to be able to differentiate between a DAESH insurgent emplacing a roadside bomb or preparing an ambush position from those civilians digging a latrine or napping in the shade of a tree.

Google's embarrassment involving its image-recognition algorithm demonstrates O'Neil's concern in multiple ways. In 2015 Google's algorithm consistently tagged black people as "gorillas." This was not the doing of some racist employee, but a result of the uncertainty of machine learning. In

---

[22] Caton, *Autonomous Weapon Systems*, 52.
[23] Scott Shane and Daisuke Wakabayashi, "'The Business of War': Google Employees Protest Work for the Pentagon," The New York Times, April 04, 2018, accessed April 10, 2018, https://www.nytimes.com/2018/04/04/technology/google-letter-ceo-pentagon-project.html.

response, Google removed the search terms to prevent the error from continuing to happen. However, over two years have elapsed without a "more sustainable fix…highlight[ing] the extent to which machine learning technology, which underpins the image recognition feature, is still maturing."[24] Image recognition must be far more reliable if it will be used to aid in distinction, especially if the technology will be used overseas where cultural differences will only serve to compound the problem. The alternative is to start curating the information like Google did, or by using unspecific terms that will introduce their own level of ambiguity, to a task that demands perfection.

The capabilities of AI, though promising, are not at a sufficient level that would assure distinction when a lethal weapon system is employed. Perfect, however, may not be prerequisite for allowable. Indeed, current technologies, if embodied in a lethal weapon system, may amount to having a half-blind sniper providing overwatch support on an operation; only slightly less likely to hit a friendly soldier than an enemy. The lethal weapon system's inability for distinction excludes it from use. This prohibition exists, however, only because friendly units or innocents are in the vicinity. In fact, Schmitt offers that such a system incapable of distinguishing between civilians and combatants "should not necessarily be unlawful per se, as autonomous weapons systems could be used in situations where no civilians were present, such as against tank formations in the desert or against warships." Executive and adaptive weapon systems may also take the fight to a different regime where only combatants operate (deep space, undersea, specific geographic locations) and may thus be used without fear of violating discrimination.

Finally, Schmitt explains that "a rifle is not prohibited under international law but using it to shoot civilians would constitute an unlawful use." Lethal weapon systems could presumably be allowed under international law so long as their use against noncombatants remains unlawful and immoral. This assumes that adaptive lethal weapon systems will fail at distinction, but

---

[24] Alex Hern, "Googles Solution to Accidental Algorithmic Racism: Ban Gorillas," The Guardian, January 12, 2018, accessed April 10, 2018, https://www.theguardian.com/technology/2018/jan/12/google-racism-ban-gorilla-black-people.

Schmitt posits that "it is categorically not the case that all such systems will do so."[25] It is conceivable that a state-of-the-art technology veiled as a state secret could exist that enables the desired effect.

Conclusion

The absence of moral agency in adaptive lethal weapon systems is not a barrier to their use. So long as the moral agent can be determined, justice may be served. For adaptive lethal weapon systems, this obligation resides with the entity whose authority is exercised; be it the operator, commander, or senior leader. Contextual elements, as always, are essential in determining culpability, so a heuristic method is not readily available. If justice is judgment of a moral agent against the moral code, however, one need only retrace the consequences set in motion to the point of decision to find the "smoking gun."

Distinction, in the same way, is more than meets the eye when adaptive lethal weapon systems are concerned. The combatant's condition, in neither Walzer's MEOS nor Orend/McMahan's interpretation, offers no significant barriers to the use of machines against warfighters - at least none that aren't already encountered by human combatants. However, the AI promise of exquisite discerning capabilities rings hollow when placed under scrutiny. Current immature capabilities cannot meet the high prerequisite to discriminate between those who may be killed and those who may not, and so may not be set loose in every regime. Still, adaptive lethal weapon systems are weapons. Like high-order munitions, there is a time and place for their judicious use.

[25] Etzioni, Pros and Cons, 258.

Chapter 6

For an Uncertain Future, an Indefinite End

*To carry out any strategy, history teaches us that **wisdom** and resources must be sufficient.*

- Jim Mattis

*The saddest aspect of life right now is that science gathers knowledge faster than society gathers **wisdom**.*

- Isaac Asimov

Moore's law and this era's technological big bang has breathed new life into fields of artificial intelligence and robotics previously drowning in the early churn of ideas and sluggish progress. Now, technology moves at an alarming pace - science fiction fueling a feverish progression underwritten by preceding success. Computers and machines do more things faster than ever before, commanding the widespread employment and trust of humans. With weapon systems now armed with incredible intelligence to match their unbelievable lethality, one must question whether humanity's automata have been awakened sufficiently to hold the power to decide who lives and who dies? Alternatively, do these machines remain solemn simulacra - unfeeling imitations - who, despite their creation in humanity's image, are incapable of any more life or liberty than a keystroke permits?

"Life" is the *alpha* and *omega* of any conversation about Just War Theory and lethal weapon systems. It is the invisible thread that pulls the concepts together; binding the value of human life with the question of whether a machine could ever be alive. Ultimately, this is why Just War Theory is so important, doubly so when applied to lethal "autonomous" weapon systems. Human life has worth. When one person unjustly takes the life of another, collectively the community demands justice - that the offender will be punished and a balance restored. Machines are known, but living machines are a mystery. These beings are alive in one's mind, animated by Hollywood imagination and fantastic stories. The images thrill because they are a work of fiction - at least until humankind's talents can make them a reality. As

technology's pace quickens from a trot to a run, it is imperative that the world ask what place independent machines have in *our* society and *our* rules. If not, humans in the future may very well be forced to ask what place they have in *machine* society and *machine* rules.

Genuinely independent or sentient robots remain on the horizon. Still, the machines of today hold measures of independence that must be reconciled with one's humanity and morality. Scientists worldwide research artificial intelligence, independent weapons, uses for big data, and other innovations and technologies that present moral and ethical concerns. Civilian applications in self-driving cars, photo/face recognition, and social-network big-data analytics appear relatively benign until a missile, gun, or bomb is joined to the technology. For better or worse, people want to know they can hold others accountable for their actions; civilian and military leaders must be prepared to offer answers and assure that the value of human life is not reduced by the perception of allowing a nonhuman to decide life or death.

The United States is a global military powerhouse and champion of democratic ideals. Like it or not - ready or not - the world watches what the nation does. Its example is the touchstone to which all others compare. This status must not be taken lightly, for if the United States is irresponsible in its use of technology or fails to uphold moral principles, it will lose influence in the world as surely as in the result of any failed military operation. Independence in lethal weapon systems must be proactively addressed to dispel misinformation and misunderstandings about the state of the world. This begins by changing the lexicon surrounding autonomy in machines. As it stands, common terminology is misleading and prevents meaningful discourse on the topic. This is seen in the conflation of similar but distinct terms and concepts like *machine autonomy* with *moral autonomy*, *autonomous* in reference to a particular operating regime with total inherent *independence*, and the *decision-making* ability of a lethal weapon system with its *compliance* with programming and optimization of options.

Accurate language is but one method that can aid better understanding and prepare warfighters, war-wagers, and their societies in the employment of lethal weapon systems. Unfortunately, the fear of killer robots, though not

entirely accurate, is palpable and must be assuaged as well. Vice Chairman of the Joint Chiefs of Staff Paul Selva is clear about his purpose and the limits of what is acceptable saying "my job as a military leader is to witness unspeakable violence on an enemy. In the end, when you send me or any soldier, sailor, airman or marine from the United States…out to defend the interests of our nation, our job is to defeat the enemy." In the execution of that charge, however, "a great deal of time is [spent] determining whether or not the tools we are developing absolve humans of the decision to inflict violence on the enemy. That is a fairly bright line that we are not willing to cross." Autonomous weapons with the capacity to deal deadly force based solely on the onboard computer and program parameters, he asserts, are unacceptable.[1]

The assurances of senior military leaders are not enough for some. Here, Aquinas provides the way to uncover important truths - *reason*. Through reasoned analysis of the point of decision and point of consequence for incidents in question one may reveal the truth - humans are further from the battlefield but are firmly tethered to the moral decision. War remains a human condition even if robots have entered the fight. Instead of agreeing with the notion that nobody can be held accountable, national leaders must recognize that justice is possible, enact policy, foster international norms, and call for conventions that explicitly name human agents as the culpable party. Until then, the potential for morally questionable decisions remains, a fog of anonymity having settled over the central issues. Leaders must dispell the haze, and in doing so, deter immoral action by demonstrating resolve to hold accountable transgressors.

The spectrum of independence reorganizes both terminology and thinking about "autonomous" platforms. The current use of the word "autonomy" has contributed to the confusion felt in the discourse relating to the morality of "autonomous" weapon systems. Current definitions allow differentiation of systems within the Department of Defense and other organizations by focusing on the level of human interaction with the system,

[1] Jim Garamone, "DoD Studies Terminator Weapons Conundrum, Selva Says," U.S. Department of Defense, August 26, 2016, accessed April 10, 2018, https://www.defense.gov/News/Article/Article/927792/dod-studies-terminator-weapons-conundrum-selva-says/.

but this is problematic in the implications that it gives to the general population who are led to believe machines are making life-and-death decisions without human input. The spectrum, however, emphasizes system independence instead. In this way, levels of automation in weapon systems can still be differentiated by defense officials, researchers, and operators, but a definite delineation is made in each category regarding inherent platform limitations.

This importance of clarifying levels of independence is readily appreciated when testing for moral agency. Moral agency is fundamental to the idea of justice, and therefore to the Just War Theory. The use of current philosophical definitions that inform the concept, however, leads to an inconclusive assessment of spectrum technologies due to the question of whether a machine can instantiate intention. Despite allowances for further examination of executive (semi-autonomous) and adaptive (autonomous) lethal weapon systems, no current technology can qualify as a moral agent. As such, Just War Theory principles - specifically *jus in bello* - apply according to a human moral agent's use of such systems, as with any other weapon. Humans remain accountable for outcomes.

Conversely, independent lethal weapon systems are judged to have the potential for moral agency. Yet even if they are accepted as having the ability for self-determining intentional choice, and humanity judges their actions against a moral standard; there is no realistic and meaningful reward or punishment that can be applied as a consequence; thus precluding these sentient systems from moral agency and the scrutiny, as legitimate actors, of the Just War Theory. The inability to hold moral agency, however, does not automatically prevent their use. Agency transference, point of decision, and point of consequence help establish an incident's chain of events to determine the cause and locus of the moral decision which set events in motion.

Even reduced to an instrument of war, the rules of *jus in bello* govern adaptive lethal weapon systems. Walzer's MEOS and Orend/McMahan's competing view of the combatant's moral condition present no issues to adaptive lethal weapon systems. Furthermore, AI capabilities, as of yet, cannot sufficiently discriminate between those who may be killed and those who may not be. This capability, thus far, is unneeded. As Work, Selva, and the whole

of the DoD assure, there is no intention to let loose lethal weapon systems to target and eliminate the enemy independently. So long as a human is in or on "the loop," the moral agent has the responsibility for discrimination - be it in the choice of a target, authorization for an attack, or the selection of an operating area.

Clausewitz affirms that "fighting, in turn, is a trial of moral and physical forces through the medium of the latter. Naturally moral strength must not be excluded, for psychological forces exert a decisive influence on the elements involved in war."[2] Just War Theory, through its guiding principles, strengthens the moral position of the just warrior. Together with a complete understanding of the implications of the use of lethal weapon systems, warfighters and war-wagers may be prepared to visit just violence upon the enemy. If new technologies like swarming or human-machine teaming that exploit fully adaptive lethal weapon systems are stigmatized as robot killers rather than tools of a deadly trade, these place limitations of their use and will unduly stress the support of the nation.

Time may reveal that adaptive lethal weapon systems are incompatible with American values and that the nation is willing to accept other risks before employing such technology. If so, let it be because of accurate and meaningful discourse of merits and failings rather than because of arguments about science-fiction fabrications. For all the amazing and frightening things machines in this age can do, the point of decision remains the dominion of humans. Technology allows one only to be farther from the execution of the commands, but one remains ever closer to the culpability of its consequence. The nation's rivals search for asymmetric advantages and recognize the promise of AI. Therefore, the nation's leaders cannot allow the peoples' concerns about the morality of lethal weapon systems to go inadequately answered. The people are the nation's strength, justice (and so Just War Theory) its shield, and technology its sword. Mattis warns "America's military has no preordained right to victory on the battlefield."[3] If preparations are not made in all three

[2] Carl Von Clausewitz et al., *On War* (Princeton, NJ: Princeton University Press, 1989), 127.
[3] Mattis, *National Defense Strategy*, 1.

regimes before battle is joined, Lady Liberty may find herself overcome by events - inside and out.

Bibliography

"Albert Einstein." History.com. Accessed April 10, 2018. https://www.history.com/topics/albert-einstein/videos/einstein-regret.

"Amazon Rekognition Demo for Defense." Amazon Web Services. November 06, 2017. Accessed April 10, 2018. https://aws.amazon.com/blogs/publicsector/amazon-rekognition-demo-for-defense/.

Anderson, Michael, and Susan Leigh Anderson. *Machine Ethics*. Cambridge University Press, 2011.

Asimov, Isaac, and Jason A. Shulman. *Isaac Asimov's Book of Science and Nature Quotations*. New York: Weidenfeld & Nicolson, 1988.

Associated Press. "Air Force Punishes 70 Airmen for Accidental Nuke-Armed B-52 Flight." Fox News. October 20, 2007. Accessed April 10, 2018. http://www.foxnews.com/story/2007/10/20/air-force-punishes-70-airmen-for-accidental-nuke-armed-b-52-flight.html.

"Autonomous." Merriam-Webster. Accessed March 20, 2018. https://www.merriam-webster.com/dictionary/autonomous.

"Autonomy." Merriam-Webster. Accessed March 20, 2018. https://www.merriam-webster.com/dictionary/autonomy.

Blakeley, Harold Whittle. *The 32d Infantry Division in World War II*. Madison, Wisc.: 32. Inf. Div. History Commission, 1957.

Britannica, The Editors of Encyclopaedia. "Moore's Law." Encyclopædia Britannica. November 29, 2017. Accessed March 20, 2018. https://www.britannica.com/technology/Moores-law.

Brown, Garrett Wallace, and Alexandra Bohm. "Introducing Jus ante Bellum as a Cosmopolitan Approach to Humanitarian Intervention." *European Journal of International Relations* 22, no. 4 (2016): 897-919. doi:10.1177/1354066115607370.

Caton, Jeffrey L. *Autonomous Weapon Systems: A Brief Survey of Developmental, Operational, Legal, and Ethical Issues*. Carlisle, PA: Strategic Studies Institute and U.S. Army War College Press, 2015.

Cebul, Daniel. "DARPA Wants to Connect Human Brains and Machines." Defense News. March 20, 2018. Accessed April 10, 2018. https://www.defensenews.com/it-networks/2018/03/19/darpa-wants-to-connect-human-brains-and-machines/.

Clausewitz, Carl Von, Michael Howard, Peter Paret, and Bernard Brodie. *On*

War. Princeton, NJ: Princeton University Press, 1989.

Coggin, Philip A. *Education for the Future: The Case for Radical Change*. Oxford: Pergamon Press, 1979.

Coletta, Damon. "Principal-Agent Theory in Complex Operations." *Small Wars & Insurgencies* 24, no. 2 (2013): 306-21. doi:10.1080/09592318.2013.778016.

Cox, Rory. "Expanding the History of the Just War: The Ethics of War in Ancient Egypt." *International Studies Quarterly* 61, no. 2 (2017): 371-84. doi:10.1093/isq/sqx009.

Crace, John, John Sutherland, and William Shakespeare. *Romeo and Juliet*. London: Doubleday, 2016.

Creighton, Jolene. "The "Father of Artificial Intelligence" Says Singularity is 30 Years Away." Futurism. February 14, 2018. Accessed March 20, 2018. https://futurism.com/father-artificial-intelligence-singularity-decades-away/.

Darayseh, Musa, and Abdelaziz Chazi. "Bank Specifics, Economics Environment, And Agency Theory: Determinants of Banking Performance In GCC." *The Journal of Developing Areas* 52, no. 4 (2018): 199-212. doi:10.1353/jda.2018.0060.

Dell'Amore, Christine. "Synthetic DNA Created, Evolves on Its Own." National Geographic. April 20, 2012. Accessed April 10, 2018. https://news.nationalgeographic.com/news/2012/04/120419-xna-synthetic-dna-evolution-genetics-life-science/.

Docherty, Bonnie Lynn. *Mind the Gap: The Lack of Accountability for Killer Robots*. New York, NY: Human Rights Watch, 2015.

Driscoll, William, Joseph Zompetti, and Susette Zompetti. *The International Criminal Court: Global Politics and the Quest for Justice*. New York, NY: International Debate Education Association, 2004.

Dubik, James M. *Just War Reconsidered: Strategy, Ethics, and Theory*. Lexington, KY: University Press of Kentucky, 2016.

Ellman, Jesse, Lisa Samp, and Gabriel Coll. *Assessing the Third Offset Strategy*. Report. CSIS International Security Program, Center for Strategic & International Studies. Vol. March 2017. Washington DC: Center for Strategic & International Studies, 2017.

Epic Games, and Microsoft Studios. Gears of War. Xbox 360 game. November 7, 2006.

Etzioni, Amitai. "Pros and Cons of Autonomous Weapons Systems (with Oren

Etzioni)." Library of Public Policy and Public Administration Happiness is the Wrong Metric, 2018, 253-63. doi:10.1007/978-3-319-69623-2_16.

"Facts - The Civil War (U.S. National Park Service)." The Civil War. Accessed April 10, 2018. https://www.nps.gov/civilwar/facts.htm'.

Finn, Ed. *What Algorithms Want: Imagination in the Age of Computing*. Cambridge, MA: MIT Press, 2017.

Frontline. "A Crime of Insanity." PBS. Accessed April 10, 2018. http://www.pbs.org/wgbh//pages/frontline/shows/crime/trial/faqs.html.

Garamone, Jim. "DoD Studies Terminator Weapons Conundrum, Selva Says." U.S. Department of Defense. August 26, 2016. Accessed April 10, 2018. https://www.defense.gov/News/Article/Article/927792/dod-studies-terminator-weapons-conundrum-selva-says/.

Gibbs, Samuel. "Elon Musk leads 116 experts calling for outright ban of killer robots." The Guardian. August 20, 2017. Accessed March 20, 2018. https://www.theguardian.com/technology/2017/aug/20/elon-musk-killer-robots-experts-outright-ban-lethal-autonomous-weapons-war.

Giles, Martin. "The GANfather: The Man Who's given Machines the Gift of Imagination." MIT Technology Review. February 27, 2018. Accessed April 10, 2018. https://www.technologyreview.com/s/610253/the-ganfather-the-man-whos-given-machines-the-gift-of-imagination/.

Grannan, Cydney. "What's the Difference Between Morality and Ethics?" Encyclopædia Britannica. Accessed April 10, 2018. https://www.britannica.com/story/whats-the-difference-between-morality-and-ethics.

Guthrie, Charles, and Michael Quinlan. *Just War: The Just War Tradition: Ethics in Modern Warfare*. London: Bloomsbury, 2007.

Ham, Paul. "As Hiroshima Smouldered, Our Atom Bomb Scientists Suffered Remorse." Newsweek. April 12, 2016. Accessed April 10, 2018. http://www.newsweek.com/hiroshima-smouldered-our-atom-bomb-scientists-suffered-remorse-360125.

Hargrove, Eugene C. *The Animal Rights, Environmental Ethics Debate: The Environmental Perspective*. Albany: State Univ. of New York Press, 1992.

Hern, Alex. "Google's Solution to Accidental Algorithmic Racism: Ban Gorillas." The Guardian. January 12, 2018. Accessed April 10, 2018. https://www.theguardian.com/technology/2018/jan/12/google-racism-ban-gorilla-black-people.

Hew, Patrick Chisan. "Artificial Moral Agents Are Infeasible with Foreseeable Technologies." *Ethics and Information Technology* 16, no. 3 (May 13, 2014): 197-206. doi:10.1007/s10676-014-9345-6.

Hijiya, James A. "The "Gita" of J. Robert Oppenheimer." *Proceedings of the American Philosophical Society* 144, no. 2 (June 2000): 123-67. Accessed April 10, 2018. http://www.jstor.org/stable/1515629.

Himma, Kenneth Einar. "Artificial Agency, Consciousness, and the Criteria for Moral Agency: What Properties Must an Artificial Agent Have to Be a Moral Agent?" *Ethics and Information Technology* 11, no. 1 (2008): 19-29. doi:10.1007/s10676-008-9167-5.

Hintze, Arend. "Understanding the Four Types of AI, From Reactive Robots to Self-Aware Beings." The Conversation. November 13, 2016. Accessed April 10, 2018. http://theconversation.com/understanding-the-four-types-of-ai-from-reactive-robots-to-self-aware-beings-67616.

"Horrors of Next War Pictured by Hobgood In Talks to Rotarians." *Greensboro Daily News*, February 24, 1932.

Hughes, Thomas P. *Rescuing Prometheus*. New York, NY: Pantheon Books, 1998.

Humphreys, Joe. "Was the 1916 Rising Morally Justified?" The Irish Times. August 22, 2014. Accessed April 10, 2018. https://www.irishtimes.com/culture/was-the-1916-rising-morally-justified-1.1902758.

I, Robot. Directed by Alex Proyas. Performed by Will Smith, Bridget Moynahan, and Bruce Greenwood. Beverly Hills, CA: 20th Century Fox, 2004. Film.

Independence Hall Association. "The Declaration of Independence: Full Text." USHistory.org. Accessed April 10, 2018. http://www.ushistory.org/declaration/document/.

"Jihād as Defense: Just-war theory in the Quran and Sunnah." Yaqeen Institute for Islamic Research. Accessed March 20, 2018. https://yaqeeninstitute.org/en/justin-parrott/jihad-as-defense-just-war-theory-in-the-quran-and-sunnah/.

"John F. Kennedy: Remarks at the Wheeling Stadium, Wheeling, West Virginia. - September 27, 1962." The American Presidency Project. Accessed April 15, 2018. http://www.presidency.ucsb.edu/ws/?pid=8908.

Johnson, Deborah G., and Keith W. Miller. "Un-making Artificial Moral Agents." *Ethics and Information Technology* 10, no. 2-3 (2008): 123-33. doi:10.1007/s10676-008-9174-6.

Johnson, James Turner. *Can Modern War Be Just?* New Haven: Yale, 1986.

Kurzweil, Ray. *The Age of Spiritual Machines: When We Will Live, Work and Think in the New Age of Intelligent Machines*. New York: Penguin, 2000.

"Landmines – UNODA." United Nations. Accessed April 10, 2018. https://www.un.org/disarmament/convarms/landmines/.

Larson, Selena. "WannaCry Cyberattacks Are Still Happening. Just Ask Honda." CNNMoney. June 21, 2017. Accessed April 10, 2018. http://money.cnn.com/2017/06/21/technology/wannacry-honda-auto-shut-down/index.html.

Lee, Roy S. *The International Criminal Court: The Making of the Rome Statute: Issues, Negotiations and Results*. The Hague: Kluwer Law International, 2002.

Matousek, Mark. "The Most Impressive Things Tesla's Cars Can Do in Autopilot." Business Insider. January 29, 2018. Accessed March 20, 2018. http://www.businessinsider.com/tesla-autopilot-functions-and-technology-2017-12#musk-says-that-in-full-self-driving-mode-teslas-will-be-at-least-twice-as-safe-as-they-would-be-with-a-human-driver-at-the-wheel-15.

Mattis, Jim. *National Defense Strategy*. Washington D.C.: Department of Defense, 2018.

Mets, David R. *RPAs: Revolution or Retrogression?* Maxwell AFB, Ala.: Air University Press, Air Force Research Institute, 2010.

Mindell, David A. *Digital Apollo: Human and Machine in Spaceflight*. Cambridge, MA, u.a.: MIT Press, 2011.

Moseley, Alexander. "Just War Theory." Internet Encyclopedia of Philosophy. Accessed March 11, 2018. http://www.iep.utm.edu/justwar/.

Mozur, Paul. "Google's AlphaGo Defeats Chinese Go Master in Win for A.I." The New York Times. May 23, 2017. Accessed March 20, 2018. https://www.nytimes.com/2017/05/23/business/google-deepmind-alphago-go-champion-defeat.html.

Murphy, James G. *War's Ends: Human Rights, International Order, and the Ethics of Peace*. Washington, DC: Georgetown University Press, 2014.

"My Lai." BBC NEWS | Asia-Pacific | Murder in the Name of War. July 20, 1998. Accessed April 15, 2018. http://news.bbc.co.uk/2/hi/asia-pacific/64344.stm.

Orend, Brian. *The Morality of War*. Peterborough (ON): Broadview Press, 2013.

Pellerin, Cheryl. "Project Maven Industry Day Pursues Artificial Intelligence for

DoD Challenges." U.S. Department of Defense. October 27, 2017. Accessed April 10, 2018. https://www.defense.gov/News/Article/Article/1356172/project-maven-industry-day-pursues-artificial-intelligence-for-dod-challenges/.

Pellerin, Cheryl. "Project Maven to Deploy Computer Algorithms to War Zone by Years End." U.S. Department of Defense. July 21, 2017. Accessed April 10, 2018. https://www.defense.gov/News/Article/Article/1254719/project-maven-to-deploy-computer-algorithms-to-war-zone-by-years-end/.

Plato. *The Republic*. John Wiley & Sons, 2012.

Prado, Guia Marie Del. "These Weapons Can Find a Target All by Themselves - And Researchers Are Terrified." Business Insider. July 30, 2015. Accessed April 10, 2018. http://www.businessinsider.com/which-artificially-intelligent-semi-autonomous-weapons-exist-2015-7.

National Science and Technology Council (U.S.). Committee on Technology. Subcommittee on Machine Learning and Artificial Intelligence, *Preparing for The Future of Artificial Intelligence*.

Price, Rob. "Microsoft Is Deleting Its AI Chatbot's Incredibly Racist Tweets." Business Insider. March 24, 2016. Accessed April 10, 2018. http://www.businessinsider.com/microsoft-deletes-racist-genocidal-tweets-from-ai-chatbot-tay-2016-3.

Ramsey, Boniface, Roland J. Teske, and Augustine, Saint, Bishop of Hippo, 354-430. *Contra Faustum Manichaeum*. Hyde Park, NY: New City Press, 2007.

Reichberg, Gregory M., Endre Begby, and Henrik Syse. *The Ethics of War: Classic and Contemporary Readings*. Malden, MA: Blackwell Publishing, 2013.

Robinson, Paul F. *Just War in Comparative Perspective*. Aldershot: Ashgate, 2003.

Rosen, Rebecca J. "'I've Created a Monster!' On the Regrets of Inventors." The Atlantic. November 23, 2011. Accessed April 10, 2018. https://www.theatlantic.com/technology/archive/2011/11/ive-created-a-monster-on-the-regrets-of-inventors/249044/.

Rosenberg, Matthew, and John Markoff. "The Pentagon's 'Terminator Conundrum': Robots That Could Kill on Their Own." The New York Times. October 25, 2016. Accessed March 20, 2018. https://www.nytimes.com/2016/10/26/us/pentagon-artificial-intelligence-terminator.html.

S, T. "How Does a Self-Driving Car Work?" The Economist. May 12, 2015.

Accessed April 10, 2018. https://www.economist.com/blogs/economist-explains/2013/04/economist-explains-how-self-driving-car-works-driverless.

Sample, Ian. "Ban on Killer Robots Urgently Needed, Say Scientists." The Guardian. November 13, 2017. Accessed March 20, 2018. https://www.theguardian.com/science/2017/nov/13/ban-on-killer-robots-urgently-needed-say-scientists.

Sample, Ian. "Organisms Created with Synthetic DNA Pave Way for Entirely New Life Forms." The Guardian. January 24, 2017. Accessed April 10, 2018. https://www.theguardian.com/science/2017/jan/23/organisms-created-with-synthetic-dna-pave-way-for-new-entirely-new-life-forms.

Schlosser, Markus. "Agency." Stanford Encyclopedia of Philosophy. August 10, 2015. Accessed April 10, 2018. https://plato.stanford.edu/entries/agency/.

Shane, Scott, and Daisuke Wakabayashi. "'The Business of War': Google Employees Protest Work for the Pentagon." The New York Times. April 04, 2018. Accessed April 10, 2018. https://www.nytimes.com/2018/04/04/technology/google-letter-ceo-pentagon-project.html.

Shanker, Thom. "2 Top Leaders of Air Force Pushed Out After Inquiry." The New York Times. June 6, 2008. Accessed April 10, 2018. https://www.nytimes.com/2008/06/06/washington/05cnd-military.html.

Simonite, Tom. "Sorry, Banning Killer Robots Just Isn't Practical." Wired. August 24, 2017. Accessed April 10, 2018. https://www.wired.com/story/sorry-banning-killer-robots-just-isnt-practical/.

Singer, Peter W. *Wired for War: The Robotics Revolution and Conflict in the Twenty-first Century*. New York: Penguin Books, 2010.

Singer, Peter W. "Isaac Asimov's Laws of Robotics Are Wrong." Brookings. July 28, 2016. Accessed April 10, 2018. https://www.brookings.edu/opinions/isaac-asimovs-laws-of-robotics-are-wrong/.

Sisk, Richard. "Mattis' Pet Peeve: Calling Drones 'Unmanned Aerial Vehicles'." Military.com. February 21, 2018. Accessed March 20, 2018. https://www.military.com/defensetech/2018/02/21/mattis-pet-peeve-calling-drones-unmanned-aerial-vehicles.html.

Skinner, B. F. *Contingencies of Reinforcement: A Theoretical Analysis*. Englewood Cliffs, NJ: Prentice Hall, 1969.

Snow, Jackie. "The US Military Wants AI to Dream Up Weird New Helicopters." MIT Technology Review. March 26, 2018. Accessed April 10, 2018. https://www.technologyreview.com/s/610454/the-us-military-wants-ai-to-dream-up-weird-new-helicopters/.

"Star Trek Borg." StarTrek.com. Accessed April 10, 2018. http://www.startrek.com/database_article/borg.

Star Wars Episode VII: The Force Awakens. Directed by J.J. Abrams. Performed by Daisy Ridley, John Boyega, and Oscar Isaac. United States: Lucasfilm, 2015. Film.

Starr, Barbara. "Air Force Investigates Mistaken Transport of Nuclear Warheads." CNN. September 6, 2007. Accessed April 10, 2018. http://www.cnn.com/2007/US/09/05/loose.nukes/index.html.

Steinbeck, John. "Of Mice and Men." Chapter 5. Accessed April 10, 2018. https://www.cliffsnotes.com/literature/o/of-mice-and-men/summary-and-analysis/chapter-5.

Strawser, Bradley Jay. *Killing by Remote Control: The Ethics of an Unmanned Military*. New York, NY: Oxford University Press, 2013.

Sullivan, Danny. "How Machine Learning Works, As Explained by Google." MarTech Today. May 03, 2017. Accessed April 10, 2018. https://martechtoday.com/how-machine-learning-works-150366.

Tan, Oliver. "How Does A Machine Learn?" Forbes. September 08, 2017. Accessed April 10, 2018. https://www.forbes.com/sites/forbestechcouncil/2017/05/02/how-does-a-machine-learn/.

Taylor, Angus MacDonald, and Angus MacDonald Taylor. *Animals and Ethics: An Overview of The Philosophical Debate*. Peterborough, Ont.: Broadview, 2003.

Saint Augustine, Bishop of Hippo. *The City of God*. Translated by Marcus Dods. Vol. 1. T. & T. Clark, 1888. Original from Columbia University, Digitized Jul 13, 2009

Thucydides, Robert B. Strassler, and Richard Crawley. *The Landmark Thucydides: A Comprehensive Guide to The Peloponnesian War*. New York: Free Press, 2008.

"The Only Thing Necessary for the Triumph of Evil Is That Good Men Do Nothing." Quote Investigator. Accessed April 10, 2018. https://quoteinvestigator.com/2010/12/04/good-men-do/.

Thompson, Hugh. *Moral Courage in Combat: The My Lai Story: Lecture*. Annapolis, MD: U.S. Naval Academy, Center for the Study of

Professional Military Ethics, 2003.

United States. Department of Defense. Office of the Secretary of Defense. *DoDD 3000.09: Autonomy in Weapon Systems*. By Ashton B. Carter.

"U.S. Accidentally Sent Ballistic Missile Parts to Taiwan." The New York Times. March 25, 2008. Accessed April 10, 2018. https://www.nytimes.com/2008/03/25/world/asia/25iht-taiwan.4.11413803.html.

U.S. Army Center of Military History. "Medal of Honor Recipients - Afghanistan." Medal of Honor Recipients - Afghanistan. March 16, 2016. Accessed April 10, 2018. https://history.army.mil/moh/afghanistan.html.

U.S. Army Training and Doctrine Command, and Army Capabilities Integration Center. *The U.S. Army Robotic and Autonomous Systems Strategy*. Fort Eustis, VA: U.S. Army Training and Doctrine Command, 2017.

UNICEF. "Land Mines: Hidden Killers." Land Mines: Hidden Killers. October 15, 1995. Accessed April 10, 2018. https://www.unicef.org/sowc96pk/hidekill.htm.

Walzer, Michael. "The Ethics of Warfare in the Jewish Tradition." *Philosophia* 40, no. 4 (2012): 633-41. doi:10.1007/s11406-012-9390-5.

Walzer, Michael. *Just and Unjust Wars A Moral Argument with Historical Illustrations*. New York, NY: Basic Books, 2015.

Watson, D. P., and D. H. Scheidt. "Autonomous Systems." *John Hopkins APL Technical Digest* 26, no. 4 (2005): 368-76.

"What Is the Cloud?" What Is the Cloud - Definition | Microsoft Azure. Accessed April 10, 2018. https://azure.microsoft.com/en-us/overview/what-is-the-cloud/.

Winnefeld, James A., Vice Chairman of the Joint Chiefs of Staff, and Frank Kendell, Under Secretary of Defense (Acquisition, Technology and Logistics). *Unmanned Systems Integrated Roadmap, FY2013-2038*. Vol. Reference Number 14-S-0553. Washington, DC: Department of Defense, 2013.

"World | Heroes of My Lai Honoured." BBC News. March 07, 1998. Accessed April 10, 2018. http://news.bbc.co.uk/2/hi/special_report/1998/03/98/mylai/62924.stm.

Yannakogeorgos, Panayotis A., and John P. Geis. *The Human Side of Cyber Conflict: Organizing, Training, And Equipping the Air Force Cyber Workforce*. Maxwell Air Force Base, Alabama: Air University Press, Air Force Research Institute, 2016.

Zetlin, Minda. "5 TED Talks on AI to Watch." The Enterprisers Project. October 20, 2017. Accessed April 10, 2018. https://enterprisersproject.com/article/2017/10/5-ted-talks-ai-watch.