Introduction

Dual Moral Responsibility and the
Ethical Challenges of Twenty-First-Century Warfare

STEVEN C. ROACH AND AMY E. ECKERT

The growing reliance on autonomous weapons systems (AWS) has spurred much debate about the ethics of warfare. In recent years, AWS has called attention to the rapidly changing dimension of modern warfare in which unmanned weapons will be able to detect targets and even attack satellite systems in space. For many, this highlights the need for more ethical thinking and theorizing about the moral dimensions and implications of AWS (see Gunkel 2018; Buhta et al. 2016; Sparrow 2012; Strawser 2013). The aim of this volume is to examine these dimensions and implications through various revisionist applications of just war theory to modern warfare, in particular AWS. In doing so, we seek to build on current applications of just war principles, including the rules of conduct in war (*jus in bello*) for AWS (Arkin 2010), and to address the moral risks and possibilities of extending just war principles to the artificial intelligence (AI) of AWS. The hope is that this will enable us, as Christian Brose (2019, 131–32) points out, "to focus more energy on making moral decisions about the intended outcomes of warfare."

Still, the challenge facing many ethicists of war, particularly just war theorists, is that AWS is a practical concept that appears to lie outside the human(ist) scope of the just war tradition. By critiquing the statist limitations of traditional and legalist just war theory and focusing attention on the individual human rights of actors (while also examining

the legitimate status of other actors in the international realm, such as private military and security companies), revisionist just war theory has largely refrained from addressing questions regarding the legal and moral status of AWS (Gruszczak 2018, 34–35). Revisionist just war theory remains largely, if not exclusively, concerned with contesting international law and state responsibility to the extent that it uses these domains to legitimize *jus ad bellum* (right to war) and to enforce the codes of moral conduct during war (McMahan 2009; Fabre 2012). Seth Lazar (2017, 41) summarizes the revisionist challenge as follows:

> The archetypical traditionalist . . . is a nonreductivist collec-tivist who uses realistic cases. The archetypical revisionist is an individualist reductionist who uses cases involving meteors and mind control. Simplifying a little, we might unify the former positions under the heading of political philosophy approaches to just war theory and group the latter together as moral philosophy approach.

This revisionist challenge problematizes many of the assumptions on which the more traditional approach rests. Todd Burkhardt, for example, argues "that the issue of fighting with the right intention also requires us to understand the conditions for a just and lasting peace" (Burkhardt 2017, 1). Just war theory's narrow focus thus reflects the trouble with relying on conventional facts of human intelligence, discretion, and emotion to justify the ethics of going to war or preserving peace. Working beyond this restrictive focus requires us to critically understand the many ways that violence and advanced weapons systems marginalize persons. Much of this will in turn depend on the ethical engagement with different actors and AWS that allows us to contest and work beyond the limits of just war theory's revisionist and conventionalist applications.

   This volume represents a critical engagement with this emerging gap(s) in logic regarding the moral responsibility of individuals, states, and AWS. As part of this engagement, it adopts the notion of dual moral responsibility, or the differing yet related notions of moral responsibility and legitimate authority, to analyze the changing roles and behaviors of various new actors in the global realm. Dual moral responsibility encapsulates the conflicts and contradictions driving the need for violent weapons and technologies to secure peace and to neutralize the effects of violence. At the same time, the changing conditions of warfare,

including the rapidly advancing technology of AWS, are challenging the way we theorize and apply just war principles. If just war theory is to address this challenge, it must begin to confront and engage the biases and conventional logic that fix or restrict its ethical content, including the ideas and assumptions (e.g., gender and individual human agent) used to theorize about just war. The trouble, in other words, with relying on conventional biases to analyze AWS is that it exposes the very limits of these biases against the ethical potential of AWS, or the moral and legal status of AWS. We therefore need to take more seriously the moral implications of the adaptive capacity of AWS (to learn from their environment)—and whether this supports the moral autonomy subsumed under just war principles such as legitimate authority. As Robert Sparrow writes, "The use of such systems may render the attribution of responsibility for the actions of AWS to their operators problematic. . . . where such use of autonomous weapons seems to risk a responsibility gap and where this gap exists, it will not be plausible to hold that when a commander sends AWS into action, he or she is acknowledging the humanity of those the machines eventually kill" (Sparrow 2015, 108). If AWS are truly autonomous, then one could argue that they operate beyond the restraint of human operators, that is, they do not rise to the level of moral personality that would qualify them to bear responsibility for their actions (Roff 2013a). The problem is that in the absence of any assigned responsibility for the actions of AWS, it becomes increasingly problematic to deploy them.

A further complication of AWS is the erosion of the line between war and peace. The availability of low-level, low-risk force means an expanding gray area between all-out war and peaceful relations between states. Michael Walzer, for instance, devised new principles to understand this gray area, *jus ad vim*, or the ethics of using force, short of war (Brunstetter and Braun 2013). But this new set of rules also raises a very real danger of perpetuating violence or legitimizing the use of low-level force outside the theater of broader conflict (Enemark 2014). Diffusing low-level force, in other words, can erode the boundaries between conflict and peace. Because of the proliferation of actors, mapping the responsibilities of nonhuman actors becomes an even more daunting task in terms of legitimizing moral authority and autonomy. Both the just war tradition and the ethics of AWS, then, describe different and often opposing ways of analyzing the effects of violence and modern warfare (i.e., new actors with increased technological precision as well as privatized control).

The contributors address this tension by reevaluating the ethical constructs of just war theory, such as moral responsibility, accountability, autonomy, and rights, and by using various empirics of artificial intelligence (AI) to formulate the ethical applications of AWS. In doing so, they take up a range of thematic issues, including the deep morality of war, the reconfiguration of war ethics, the possible end of just war, a moral groundwork for robot rights and responsibilities, and the ethical uncertainty of advancing morality and accountability (i.e., compliance with the laws of war).

Just or Unjust Warfare

Just war tradition as we know it can be traced to the religious writings of Augustine, the fourth-century Catholic saint (Brunstetter and O'Driscoll 2017). Its emergence in the writings of Francisco de Vittoria, who opposed the Spanish conquest of the new world in the sixteenth century, stressed the importance of moral and personal virtue, particularly the restraint embodied in religious devotion and piety. However, in the early seventeenth century, the focus on virtue gave way to the application of legal principles in war, or a legalist approach to just war theory. Hugo Grotius, a seventeenth-century Dutch jurist, who was the most prominent proponent of this legalist approach, focused on secular international humanitarian law and its principles governing decision-making and conduct in war.

This legal framework was dramatically expanded in the nineteenth and twentieth centuries and codified in the 1949 Geneva Conventions, which stipulate the rules and procedures for treating noncombatants humanely during war. *Jus ad bellum* (right to war), *jus in bello* (rules of conduct), and *jus post bellum* (justice after war) constitute the basis of broad criteria for justifying the conduct of warring parties and administering justice after war, including just intention, legitimate cause—usually in the form of self-defense—the probability of success, legitimate/competent authority, and proportionality or the use of force that does not rise above the level of threat. Legitimate cause and proportionality serve as the morally objective guideposts for humanitarian military interventions, providing key normative constraints on harming/injuring civilians or noncombatants through the neutrality of the rules of war and state consent (see Lang, O'Driscoll, and Williams 2013).

As Dan Caldwell and Robert Williams (2006) argue, upholding these constraints is not sufficient to determine how peace will be upheld *jus post ad bellum*. The moral reality of war may be such that even proportionality and success of the war fail to justify authority, as we have seen with the insurgencies in Libya and Iraq. If anything, this seems to reflect the growing divide between a legalist approach, or the focus on international law and the state to justify *jus ad bellum*, and a revisionist perspective on just war theory, which, in critiquing traditionalism, emphasizes the individual in terms of justifying war through a uniform code of moral conduct (McMahan 2009). In short, the legalist tradition stresses how laws of war have evolved through international customary law that has imposed (legal) constraints and duties on states, including necessity and the right of self-defense.

In *Just and Unjust Wars*, Michael Walzer (2009) argues that the rules regulating the decision to go to war on humanitarian grounds were constituted by a long-standing moral tradition comprising the opinions of political leaders, philosophers, and scholars. Yet for Walzer, the moral reality of suffering never corresponded to the political will to redress such suffering. Rather, the virtues of the early legal tradition (in objectifying the rationale for just war) lost sight of the political reality shaping these principles and the attendant responsibility to treat such suffering as a condition for invoking these legal principles. In his view, such reality reflected how moral actions redounded to the political advantage of coercive powers and the values they sought to project overseas.[1] But while the mass killing of civilians may constitute a legitimate cause of armed intervention, it does not ensure that either a competent or trustworthy authority will enforce just war norms. In many respects, enforcing such norms through the UN Security Council—where the veto power of the permanent member states can rescind these morally compelling cases as we have seen with Syria—can dramatically raise the political stakes of promoting peace and security or of countering terrorism. Relying on international law to justify the right of states to defend themselves—and hence to go to war via the Law of Armed Conflict—assumes a difficult trade-off between state constraints and the moral failure to stop mass killings (Morkevicius 2018). This stems from the fact that states remain the primary actors or the sole source of moral responsibility, rather than the individual victims of violence, whose rights and protections have become subordinated to the moral responsibility of states.

Thus, the trouble with relying on state compliance to legitimize the responsibility to go to war is that we tend to downplay the importance of international human rights law. And yet in recent years, we have seen how gross violations of international human rights have justified recent decisions to launch humanitarian wars in Kosovo and Libya. Why should only states, then, be treated as the sole agents for justifying the right to wage war and validating the rules and moral conduct of the Law of Armed Conflict? In addition, how does the legalist focus expose the need for focusing more on human virtue to determine the nature of just cause? The idea is that understanding basic human rights protection involves more than simply the application of state rules, norms, and constraints against killing combatants and noncombatants in war: it also concerns the moral status of individuals.

Revisionist just theory adopts this idea to analyze the gap in logic between the legalist approach and an individual rights–based approach via the evolving codes of moral conduct (e.g., the Law of Armed Conflict). It therefore focuses on the changing conditions of individual moral responsibility, moral intention, and legitimate authority and argues that the Law of Armed Conflict should abide by international human rights law (see Brunstetter and Holeindre 2018; Burkhardt 2017; Eckert 2016; Gentry and Eckert 2014). Revisionist just war theory is also predicated on a cosmopolitan perspective of the individual, that is, the moral equality of individuals. As Cecile Fabre (2012, 7) argues, "All individuals, wherever they reside, have the aforementioned rights against everyone else, irrespective of residence." Global justice in this sense points to the gap between revisionists and traditionalists, in which revisionists reject the need for any authority beyond the individual, even though they do not deny the political reality of states and other communities.

One of the strategies used by revisionists is to stress the role played by compassion and empathy in reinforcing just war norms and furthering moral commitments. But hostile emotions can also fuel the impulsive actions that pressure and destabilize these norms. In addition, the conflicts between nonhostile and hostile emotions often compel us to dichotomize between compassion (or positive emotion) and hatred (negative emotion) and to rely on restraint to control the effects of hostile emotions (see Solomon 2007). Hostile emotions, it can be said, constitute the irrational elements of moral judgment. By the same token, positive emotions can exceed the ability to control the effects of humanitarian wars. In the Kosovo War, for instance, while upholding the basic human rights of

Kosovar civilians against the attack of states helped fuel the moral and humanitarian war, it also exposed the disparity between political and military objectives on the ground. In effect, the necessity of protecting the needs and rights of those suffering led to military measures, such as surgical strikes, that failed in most instances to support the political objectives of the war (i.e., forcing Slobodan Milosevic to surrender). Although revisionist theory enables us to understand the moral responsibility to launch wars on humanitarian grounds, it also constitutes an anthropocentric approach that has little to say about the emotive possibilities of robots or AWS.

Unsurprisingly, much of the extended focus of *jus in bello* to AWS has come from ethical scholars of AWS, such as Ronald Arkin and, to lesser extent, Robert Sparrow. For them, as we shall explain, AWS represents the possibility of eliminating the confusion and violent side effects arising from human emotions. Arkin, in fact, claims that the ethical autonomy of robots would effectively rule out the complicating influence of emotions (e.g., anger and fear). As such, the increasing presence of robotic weapons reinforces the need for reconciling these two ethical approaches and calls attention to how the increasing presence of AWS will require deeper ethical inquiry into the nonhuman sources of action and conduct. Thus, as Peter Singer (2009, 390) puts it, "By replacing human judgment with AI technologies, it becomes possible to limit the effects of war." Because artificial intelligence introduces precision weaponry, it raises the possibility of eliminating much of the violence caused by human error. The laws of war could in this manner tap into the advanced capacity of robots to carry out orders. For Arkin, AWS provide the opportunities for more precise compliance with the rules of lawful combat and just war criteria as stipulated in the Geneva Conventions.

In this way, Arkin stresses a dualistic notion of moral responsibility of human and nonhuman agents and restricts the emotive possibilities of the legal and moral status of robots' adaptive capacity. Yet it is this amoral efficacy of emotion, or absence of emotion, that remains in tension with a revisionist just war focus and that raises the difficult question of how best to use (an evolving notion) the notion of dual moral responsibility to work beyond the assumptions of robots' legal status, especially concerning the efforts to confront the legal gap between AWS and moral accountability. Indeed, as David Gunkel (2018) claims, insisting that robots cannot think also constitutes an event that precludes the normative possibility of robots' rights. Any strong and deepening

engagement with these dualistic aspects of moral responsibility, then, needs to contend with this possibility.

The Ethical and Moral Challenges
of High-Tech Warfare

With more and more military robots being programmed to fight and, in most cases, kill, the lethal use of AWS can be said to constitute an emerging reality in warfare. This also represents the practical opportunity to address some of the more vexing issues of indiscriminate violence and killing that have seemed to haunt conventional just war theorizing (i.e., the harm done to innocent civilians). For instance, if robots will be able to finally target combatants with greater precision, then it might be possible to avert the consequences and effects of the fog of war, such as the indiscriminate killing of noncombatants or the collateral damage of drone warfare (Enemark 2015; Kaag and Kreps 2014). This suggests that somehow such weapons may be reliably programmed to discriminate between combatants and noncombatants, which in turn would mean accepting and promoting the autonomous function of weapons systems. To borrow Sparrow's definition of an autonomous weapon, such a weapon is one "capable of being tasked with identifying possible targets and choosing which to attack without human oversight and that is sufficiently complex such that, even when it is functioning perfectly, there remains some uncertainty about which objects and/or persons, it will attack and why" (Sparrow 2015, 95).

One might argue, then, that there are no assurances that autonomous weapons, with highly complex sensors and satellite tracking devices that allow the weapons system to function with almost complete certainty, will be able to perform safely without human supervision. Even the fully autonomous self-driving car cannot entirely replicate the human ability to decipher the obstacles ahead. The difference between the self-driving sensors and human sensory experience reflects what some have called "split responsibility" (which unlike, dual responsibility, conveys unrelated sources/triggers of action), in which the human driver reacts and acts differently to avoid obstacles. In terms of lethal autonomous weapons systems (LAWS), the problem has produced similar concerns and has raised the political stakes for resolving the indiscriminate killing of noncombatants. As Singer (2009, 403) points out, "Robots have great

difficulty interpreting context, and, at least, until they match humans in intelligence, it simply does not make sense to interpret a machine as having the equivalent of human rights of self-defense." There is also the (dark) possibility of robots using their developing intelligence to target humans, a fear that has been the subject of various science fiction films, such as the *Terminator* and *Bladerunner* series, and that only seems to project existing biases against robotic intelligence (i.e., that they are amoral agents). Despite this scenario, however, some countries remain resigned to develop autonomous weapons, even though many remain wary about their use in warfare.

Still, several powerful countries' military strategists are betting on this idea: namely, that autonomous robotic weapons will offer them an advantage militarily. In fact, the United States., Britain, and China have already begun research on the development of new lethal autonomous weapons systems, or advanced robotic weapons systems that carry their own sense detectors and are considered in this sense to be semiautonomous weapons (Topol 2016). In 2015, the United States, for instance, unveiled the design and section of the X-47B, a new pod-shaped aircraft that can be autonomously refueled in midair, while Britain, not to be outdone, is working on the Tauris aircraft equipped with automatic laser sensors. With nearly USD 72 billion invested in such technology, the United States continues to maintain that such advanced technology poses few risks to civilians and allows the US to better protect itself from outside threats.

The fear, then, is that AI technologies may lead to increased (human) collateral damage owing to software malfunction or programming errors. Signs of this threat have surfaced in earlier incidents involving the limited supervision of LAWS, or semiautonomous LAWS. In 2007, a South African semiautonomous antiaircraft system accidently fired upon and killed seven South African soldiers; and in 1988, the US air defense system mistakenly shot down an Iranian passenger jet (Gubrud 2016). The thorny issue is whether killer robots can be held to account for their actions; for with no human at the helm, it becomes increasingly unclear as to how to prosecute the destructive actions of robots. The only real option may be to file civil charges, effectively holding the civil programmers of these robots liable for damages. But this is not likely to curb the destructive actions of what some have called killer robots, since it will involve a high burden of proof, or depend on whether the maker had knowledge of a programming defect.

The problem with bridging this gap between international criminal law and LAWS is that AI adaptation currently remains undeveloped in relation to the rules and procedures for assigning moral responsibility. This leaves open the question of whether such responsibility can be assigned when the targeted individual lacks any sense of guilt or conscious intent (see Marchant et al. 2011). Will there be different and fair standards established for human and nonhumans, for instance? And will moral punishment for war crimes remain grounded in command responsibility (of the individual programing the robots) and, as such, displace the guilt and intent of a growing population of killer robots within the corpus of international criminal law? If there are two independently evolving tracks of human and nonhuman warfare, will the latter require a whole new conception of rights and autonomy to generate the efficacy of international criminal norms and the many new rules of procedure for determining the intent and knowledge of war criminals? The International Criminal Court and International Criminal Tribunals, for example, have brought hundreds of war criminals to justice and arguably helped deter criminal behavior (Roach 2006, 2013). But such a deterrent effect, which relies on the capacity of courts to expose the knowledge of perpetrators' intent, cannot, at least for now, apply to autonomous robots programmed to kill.

Nonetheless, many lethal weapons systems such as the PHALANX Close-in System, which is a rapid-fire computer-controlled radar gun invented by General Dynamics, are only minimally capable of firing on their own or without human guidance. Military officials are still dubious about the need to remove these and more complex semiautonomous weapons (Tauris aircraft) from human decision-making. This seems to ensure that we will continue to struggle to define the nascent, open-ended parameters of dual moral responsibility. Moreover, it might also explain the polarizing responses LAWS.

For the most part, the response to this legal and moral challenge of AI has been twofold: (1) either reject or resist LAWS or (2) devise whole new ways of rethinking their evolving role (and agency) in warfare (O'Connell 2013). The former is highlighted by concerted calls for a complete ban on LAWS by Human Rights Watch (HRW) and the Campaign to Stop Killer Robots, a coalition of nongovernmental organizations (NGOs) working to ban fully autonomous weapons (which has been in the forefront of this movement to ban all LAWS) (Campaign to Stop Killer Robots 2018). In a report issued in April 2015, HRW documented the rapid rise of many semiautonomous weapons, arguing that regulation will do little to stop the destructive impact of fully

autonomous killer robots (Human Rights Watch 2015). HRW lawyers and activists recently voiced their concerns at a delegate meeting of the Convention on Certain Conventional Weapons, an agreement signed by 125 countries that has pledged to eliminate weapons that indiscriminately kill civilians (UNOG 1980).

Despite these good intentions, such resistance has done little to alter the political reality of LAWS or the most powerful countries' commitment to produce more sophisticated LAWS. As both Peter Singer and August Cole (2016) argue, it is perhaps more realistic to transition to or erect new laws and rules to hold humans accountable for any lethal mistake made by the robots. By clarifying which maker is and is not responsible, the hope is that authorities will adopt rules constraining the reckless behavior of states and corporations (Singer 2009, 20). This may be the first uncomfortable step of working toward developing and engaging the normative possibilities of their agency (or autonomy) and bridging the gap between the just war norms of legitimizing moral conduct (or just intent) and AWS. It may also be why many have sought to address the complexity of this challenge through disciplinary approaches or a multi-disciplinary approach that can map an evolutionary pathway of robotic intelligence in which intention, conscience, and even emotion (feelings) might justify a new conception of (moral) responsibility and new rights (Tripod and Wolfendale 2012; Kahn 2002; Benjamin 2013; Sparrow 2012).

In short, new ethical guidelines will be needed to regulate the moral conduct of robots and to bridge the ethical gap between the Laws of Warfare and robot intelligence. Moreover, there needs to be a larger effort to work beyond existing conventions and customs of war as well as understanding the critical ties between just war theorizing and the moral autonomy of robots.

Overview of the Book

The first part of the volume focuses on the limits and problematic aspects of just war theory and the attempts to revise and contest these limits. The most compelling set of challenges to conventional just war theory involves a "deep morality of war" approach to the ethics of killing in war, which draws from cosmopolitan political theory to question the justice of a state-based system and status-quo rules of war. The contributors to this first part of the volume seek, in creative ways, to refigure the norms of moral responsibility of just war theory.

In chapter 1, Peter Sutch addresses several developments in international law and in military technology and practice that have sparked large-scale criticisms of traditional just war theory. He provides a novel defense of legalist or conventionalist just war theory against the attacks by cosmopolitan critics. Such criticisms generate *jus ad bellum* arguments for an expanded right to war for humanitarian and defensive purposes and *jus in bello* arguments that deny the moral equality of combatants that underpins the distinction between combatant and civilian. These arguments challenge the basic principles of just war theory and the laws of war in the name of a deeper and more refined moral philosophy. They also create deep practical challenges for the humane governance of conflict. The chapter focuses on two rather different modern developments—the increasing normative importance of human rights and the evolution of military technology that enables "riskless" combat—to show the differences between legalist just war theory and that of its critics. It argues that a conventionalist understanding of human rights is both more relevant to modern warfare and still an effective critical tool of managing conflict.

The detachment from the reality of violence, then, is symptomatic of the dichotomous thinking on war and peace and violence and gender. In chapter 2, Laura Sjoberg analyzes what she calls the trichotomy: namely, unjust war (the ultimate evil); just war (necessary evil and morally permissible); and just peace (that toward which just war strives). The chapter examines various assumptions of this trichotomy, including bad violence, acceptable/good violence, and nonviolence, and argues that war and peace are conceptually and empirically problematic. Building on feminist theorizing about the links between sexism, patriarchy, and violence, Sjoberg proposes both that violence is a continuum rather than a delineable entity, and that there is no nonviolent alternative to violence. After laying out its theoretical approach to violence, she turns to exploring that interpretation's implications for just war theorizing. Here, she contends that no additive or multiplicative approach to *jus ad bellum*, *jus in bello*, or *jus post bellum* can account for thinking about violence as a continuum. A continuum approach to violence, she concludes, has a number of important implications for many just war concepts, as well as for the overall utility of just war thinking.

In chapter 3, Thomas Doyle takes up this nonconventional logic of just warfare by addressing the issue of whether states' reliance on automated weapons systems in contemporary warfare motivates a similar line of inquiry into the moral responsibility related to contemporary nuclear deterrence. In his chapter, he focuses on the extent to which state practices

of nuclear deterrence and their plans for responding to nuclear deterrence failure have generally entailed a sufficient loss of human autonomy by state leaders. To address this question, he undertakes a brief conceptual analysis of "autonomy" and the conditions under which it might be ceded by reliance on diverse modes of instrumental rationality. He then applies this analysis to the logic of the practice of nuclear deterrence, noting the strategic and moral problems that emerge with respect to the human control of nuclear force. Here he concludes that nuclear deterrence counts as a borderline case of "dual moral responsibility" insofar as it seems to compel at some level ethical detachment (but not in the sense of ethical neutrality) and a partial delegation of moral decision-making to "automatic" processes *while at the same time* retaining some measure of human control over the prospective uses of nuclear weapons. Under conditions of deterrence failure, the logic of nuclear reprisal suggests moral nihilism, and this, too, is in need of further theorization regarding moral responsibility in the case that nuclear weapons truly do, as Michael Walzer famously put it, "explode just war theory."

Another way of contesting the traditional or conventionalist logic of just war theory is through the legitimate role of unconventional actors that are challenging legalist just war theory. In recent years, private military and security companies have adopted strategies of seeking legitimacy under international law through the antimercenary norm. What this suggests is that just war criteria such as competent authority are being reconstituted by high-tech modern warfare. In chapter 4, Sommer Mitchell shows how the antimercenary norm has prohibited private actors from participating in conflict. Private military and security companies (PMSCs) have been hired regularly by state actors as well as international organizations to provide support services for military and security operations. This need to outsource warfare to PMSCs, she shows, reveals a complex struggle to acquire legitimacy through compliance with international, antimercenary norms. PMSCs, and not simply states, have begun to shape and reconstitute the (dispersed) meaning of legitimate authority in twenty-first-century warfare by altering perceptions of their commitment to human rights protections. The issue this raises is whether new human agents can help bridge the gap between conventional just war theory and robotic intelligence, and if this gap symptomizes the misguided faith in human fallibility of revisionist just war theorizing.

Part II consists of essays that seek to varying degrees to build on research supporting the moral responsibility of AWS or robots. In chapter 5, David Gunkel addresses the issue of whether the growing recognition

of the legality of robots should make way for stronger engagement with robot rights. He makes a case for this apparently marginal set of concerns by responding to a seemingly simple and direct question: Can or should killer robots have rights? This question is not just any question. Indeed, we should be clear about the inherent difficulty of even articulating such a query, since the very concept of robots having rights strains common sense or good scientific reasoning. That it needs to be purposefully avoided as something that must not be thought, insofar as it is a kind of prohibited idea or blasphemy that would open a Pandora's box of problems and therefore should be suppressed or repressed. Gunkel argues that the existing classification schema—one that recognizes only two kinds of entities, personnel or equipment—may be too restrictive and insensitive to respond to and take responsibility for the different kinds of things with which we interact and involve ourselves. Whatever the reason(s), there is something of a deliberate decision and concerted effort not to think—or at least not to take as a serious matter for thinking—the question of robot rights. In this, Gunkel explores various reasons for considering the rights of robots in general and the rights of battlefield robots.

In chapter 6, Jai Galliott analyzes the problem of how to assign moral responsibility when large groups of people, organized or unorganized, wrongfully cause some harm that is pervasive in our world given the ubiquitous nature of large organizations, such as corporations, nations, and universities, that are involved in the development and deployment of emerging military technologies. He argues that advocates of a ban on lethal autonomous systems have erroneously attempted to take the problem of many hands one step further in suggesting that said weapons systems have or will lead us to a problem of "no hands." Here he deconstructs such arguments and, in response, recharacterizes the matter confronting lethal autonomous systems as a traditional problem of many hands when traditionally conceived as the occurrence of a situation in which the collective can reasonably be held morally responsible for an outcome, even though none of the individuals can reasonably be held morally responsible for that outcome. His aim, then, is to develop a conceptual framework for moral responsibility in cases where the problem of many hands arises in the context of the design, development, or deployment of lethal autonomous weapons that facilitates the formulation and implementation of solutions.

Chapter 7 is an attempt to test and empirically support the instrumental and constitutive features of artificial intelligence as it relates

to the ethics of lethal autonomous weapons. It contends that military operations should be immune from the progress of automation and artificial intelligence evident in other areas of society. Luminary figures in science and industry as well as organized protest groups have called for an international ban on "killer robots" and the "weaponization of AI." A foundation of the argument is that to comply with international humanitarian law, autonomous weapons would need "human qualities," which, the authors argue, machines inherently lack. In contrast, the development and deployment of AI in weapons is an *ethical imperative*. A simple illustration is a weapon capable of recognizing the unexpected presence of the international protection symbol of the Red Cross in a defined target area and aborting an otherwise unrestrained human-ordered strike. This is an example of what can be called an "ethical weapon," which need not possess every human-like quality to produce a useful ethical enhancement. Ethical weapons technology is proposed to be fully integrated into the military enterprise with human commanders. The chapter outlines a case for ethical weapons and proposes a code adapted from the German Ministry of Transport and the Digital Infrastructure Ethics Commission's code for driverless automobiles. Aspects of the technological feasibility of realizing this ethical code is considered in terms of human and weapon competency, authority, and responsibility. In providing the formal semantic definitions of these concepts for encoding into weapons to provide the "meaningful human control" hitherto claimed as lacking by advocates of a ban, they seek to demonstrate the feasibility of implementing limited "human qualities" in a weapon, as part of an enterprise spanning humans and machines, in order to improve ethical outcomes.

Note

1. Thomas Hobbes's idea was that coercion was necessary to validate the contract(s) among warring parties in a state of nature.

References

Arkin, Ronald C. 2010. "The Case for Ethical Autonomy in Unmanned Systems." *Journal of Military Ethics* 9(4): 332–41.

Benjamin, Medea. 2013. *Drone Warfare: Killing by Remote Control*. London: Verso, 2013.

Brose, Christian. 2019. "The New Revolution in Military Affairs." *Foreign Affairs* 98(3): 122–34.

Brunstetter, Daniel, and Megan Braun. 2013. "From Jus ad Bellum to Jus ad Vim: Recalibrating Our Understanding of the Moral Use of Force." *Ethics & International Affairs* 27(1): 87–106. doi:10.1017/S0892679412000792.

Brunstetter, Daniel R., and John-Vincent Holeindre. 2018. The Ethics of Peace and War Revisited. Washington, DC: Georgetown University Press.

Brunstetter, Daniel R., and Cian O'Driscoll. 2017. *Just War Thinkers: From Cicero to the 21st Century*. London: Taylor and Francis.

Buhta, Nehal, Susanne Beck, Robin Geiß, Hin-Lan Yiu, and Claus Kreß. 2016. *Autonomous Weapons Systems: Law, Ethics, Policy*. Cambridge: Cambridge University Press.

Burkhardt, Todd. 2017. *Just War and Human Rights: Fighting with Right Intention*. Albany: State University of New York Press.

Caldwell, Dan, and Robert E. Williams. 2006. "Jus Post Bellum, Just War Theory and the Principles of Just Peace." *International Studies Perspectives* 7(1): 309–20.

Campaign to Stop Killer Robots. 2018. "Convergence on Retaining Control on Weapons Systems." https://www.stopkillerrobots.org/.

Curtin, Audrey Kurth. 2013. "Why Drones Fail." *Foreign Affairs* 95(5).

Eckert, Amy E. 2016. *Outsourcing War: The Just War Tradition in Age of Military Privatization*. Ithaca, NY: Cornell University Press.

Enemark, Christian. 2014. "Drones, Risk, and Perpetual Force." *Ethics & International Affairs* 28(3): 365–81.

Enemark, Christian. 2015. *Armed Drones and the Ethics of War*. New York: Routledge.

Fabre, Cecile. 2012. *Cosmopolitan War*. Oxford: Oxford University Press.

Gentry, Caron E., and Amy E. Eckert. 2014. *The Future of Just War: New Critical Essays*. Athens: University of Georgia Press.

Gruszczak, Artur. 2018. "Violence Reconsidered: Towards Postmodern Warfare." In *Technology, Ethics, and the Protocols of Modern Warfare*, edited by Artur Gruszczak and Pawel Frankowski, 26–40. New York: Routledge.

Gubrud, Mike. 2016. "Why Should We Ban Autonomous Weapons? To Survive." *IEEE Spectrum*, June 1. https://spectrum.ieee.org/automaton/robotics/military-robots/why-should-we-ban-autonomous-weapons-to-survive.

Gunkel, David. 2018. *Robots' Rights*. Cambridge, MA: MIT Press.

Human Rights Watch. 2015. "Mind the Gap: The Lack of Accountability for Killer Robots." April 9. https://www.hrw.org/report/2015/04/09/mind-gap/lack-accountability-killer-robots.

Kaag, John, and Sarah Kreps. 2014. *Drone Warfare*. London: Polity.

Kahn, Paul. 2002. "The Paradox of Riskless War." *Philosophy and Public Policy Quarterly* 22(3): 2–8.

Lang, Anthony F., Cian O'Driscoll, and John Williams. 2013. *Just War: Authority, Tradition, and Practice*. Washington DC: Georgetown University Press.

Lazar, Seth. 2017. "Just War Theory: Traditionalists versus Revisionists." *Annual Review of Political Science* 20: 37–54.

Lucas, George J. 2013. "Jus in Silico: Moral Restrictions on the Use of Cyber Warfare." In *Routledge Handbook of Ethics and War: Just War Theory in the 21st Century*, edited by Fritz Allhoff, Nicholas G. Evans, and Adam Henseke. London: Routledge.

Marchant, Gary E., Braden Allenby, Ronald Arkin, Edward T. Barret, Jason Borenstein, Lyn M. Gaudet, Orde Kittrie, Patrick Lin, George R. Lucas, Richard O' Meara, and Jared Silberman. 2011. "Institutional Governance of Autonomous Military Robots." 12 *Columbia Journal of Science and Technology Law Review* 272.

McMahan, Jeff. 2009. *Killing in War*. Oxford: Oxford University Press.

Morkevicius, Valerie. 2018. *Realist Ethics: Just War Traditions as Power Politics*. Cambridge: Cambridge University Press.

O'Connell, Mary Ellen. 2013. "Banning Autonomous Weapons." In *The American Way of Bombing*, edited by Matthew Evangelista and Henry Shue. Ithaca, NY: Cornell University Press.

Roach, Steven C. 2006. *Politicizing the International Criminal Court: The Convergence of Politics, Ethics, and Law*. Lanham, MD: Rowman & Littlefield.

Roach, Steven C. 2013. "How Political Is the ICC? Pressing Challenges and the Need for Diplomatic Efficacy." *Global Governance* 19(4): 507–23.

Roff, Heather M. 2013a. "The Strategic Robot Problem: Lethal Autonomous Weapons in War." *Journal of Military Ethics* 13(3): 211–27.

Roff, H. M. 2013b. "Responsibility, Liability, and Lethal Autonomous Robots." In *Routledge Handbook of Ethics and War: Just War Theory in the 21st Century*, edited by Fritz Allhoff, Nicholas G. Evans, and Adam Henseke, 352–64. London: Routledge.

Singer, P. W. 2009. *Wired for War: The Robotics Revolution and Conflict in the 21st Century*. New York: Penguin Books.

Singer, P. W., and August Cole. 2016. "Humans Can't Control Killer Robots, but Humans Can Be Held Accountable for Killer Robots." *ViceNews*, April 15. https://www.vice.com/en_us/article/zm75ae/killer-robots-autonomous-weapons-systems-and-accountability.

Solomon, Robert C. 2007. *True to Our Feelings: What Our Emotions Are Really Telling Us*. Oxford: Oxford University Press.

Sparrow, Robert. 2012. "War without Virtue." In *Killing by Remote Control: The Ethics of an Unmanned Military*, edited by Bradley Jay Strawser. Oxford: Oxford University Press.

Sparrow, Robert. 2015. "Robots and Respect: Assessing the Case against Autonomous Weapons Systems." *Ethics and International Affairs* 30(1): 93–116.

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

Topol, Sarah. 2016. "Attack of the Killer Robots." *Buzzfeed*, August 26. https://www.buzzfeed.com/sarahatopol/how-to-save-mankind-from-the-new-breed-of-killer-robots?utm_term=.nhmMmknd1#.hnkBKe6oN.

Tripod, Paulo, and Jessica Wolfendale, eds. 2012. *New Wars and New Soldiers: Military Ethics in the Contemporary World*. London: Routledge.

United Nations Office in Geneva (UNOG). 1980. "Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons." Geneva, October 10. https://treaties.un.org/pages/ViewDetails.aspx?src=TREATY&mtdsg_no=XXVI-2&chapter=26&clang=_en.

Walzer, Michael. 2009. *Just and Unjust Wars*. 4th ed. New York: Basic Books.

Zenko, Micha. 2016. "Do Not Believe the US Government's Official Numbers on Drone Strike Civilian Casualties." *Foreign Policy*, July 5. http://foreignpolicy.com/2016/07/05/do-not-believe-the-u-s-governments-official-numbers-on-drone-strike-civilian-casualties/.