

# The Ethics & Morality of Robotic Warfare: Assessing the Debate over Autonomous Weapons

*Michael C. Horowitz*

*Abstract: There is growing concern in some quarters that the drones used by the United States and others represent precursors to the further automation of military force through the use of lethal autonomous weapon systems (LAWS). These weapons, though they do not generally exist today, have already been the subject of multiple discussions at the United Nations. Do autonomous weapons raise unique ethical questions for warfare, with implications for just war theory? This essay describes and assesses the ongoing debate, focusing on the ethical implications of whether autonomous weapons can operate effectively, whether human accountability and responsibility for autonomous weapon systems are possible, and whether delegating life and death decisions to machines inherently undermines human dignity. The concept of LAWS is extremely broad and this essay considers LAWS in three categories: munition, platforms, and operational systems.*

MICHAEL C. HOROWITZ is Associate Professor of Political Science at the University of Pennsylvania and Associate Director of Penn's Perry World House. He formerly worked for the U.S. Department of Defense. His publications include *Why Leaders Fight* (2015) and *The Diffusion of Military Power: Causes and Consequences for International Politics* (2010). You can follow him on Twitter @mchorowitz.

The growing use of drones on today's battlefields raises important questions about targeting and the threshold for using military force. Over ninety militaries and nonstate actors have drones of some kind and almost a dozen of these have armed drones. In 2015, Pakistan shot down an Indian drone in the disputed Kashmir region, Turkey shot down a drone near its border with Syria, and both Nigeria and Pakistan acquired armed drones.<sup>1</sup>

The use of drones by the United States and others has led to an array of questions about the appropriateness of so-called remote-controlled warfare. Yet on the horizon is something that many fear even more: the rise of lethal autonomous weapon systems (LAWS).<sup>2</sup> At the 2016 Convention on Certain Conventional Weapons in Geneva, over one hundred countries and nongovernmental organizations (NGOs) spent a week discussing the potential development and use of autonomous weapon sys-

tems. An NGO, The Future of Life Institute, broke into the public consciousness in 2015 with a call, signed by luminaries Elon Musk and Stephen Hawking, as well as scientists around the world, to prohibit the creation of autonomous weapons.<sup>3</sup>

Two essential questions underlie the debate about autonomous weapons: first, would autonomous weapons be more or less effective than nonautonomous weapon systems? Second, does the nature of autonomous weapons raise ethical and/or moral considerations that either recommend their development or justify their prohibition? Ultimately, the unique facet distinguishing LAWS from non-LAWS is that the weapon system, not a person, selects and engages targets. Therefore, it's critical to consider whether the use of LAWS could comply broadly with the protection of life in war, a core ethical responsibility for the use of force; whether LAWS can be used in ways that guarantee accountability and responsibility for the use of force; and whether there is something about machines selecting and engaging targets that makes them ethically problematic. Given the centrality of these issues in debates about just war theory, it therefore makes the issue of LAWS relevant to just war theory as well.

This essay examines the potentially unique ethical and moral issues surrounding LAWS, as opposed to nonautonomous weapon systems, especially as they relate to just war theory, in an attempt to lay out some of the key topics for thinking about LAWS moving forward. It does not engage, however, with certain legal arguments surrounding LAWS, such as whether international humanitarian law implies that humans must make every individual life-or-death decision, or whether LAWS violate the Martens Clause of the Hague Convention by violating the dictates of the human conscience.<sup>4</sup> Moreover, different opponents of LAWS make different arguments, as do different critics of those opponents, so there

are undoubtedly subcomponents of each issue not discussed here. Most generally, this essay finds that the ethical challenges associated with autonomous weapons may vary significantly depending on the type of weapon. LAWS could fall into three categories: munition, platforms, and operational systems. While concerns may be overstated for LAWS that will be most akin to next-generation munitions, when thinking about autonomous weapon platforms or operational systems for managing wars, LAWS raise more important questions. Caution and a realistic focus on maintaining the centrality of the human in decisions about war will be critical.

Given the use of drones by the United States and others against terrorists and insurgents around the world, there is a tendency to conflate the entire category of military robotics with specific cases of drone strikes. However, it is a mistake to focus solely on the drone strike trees and miss the vast military robotics forest. For example, as current platforms, like the RQ-4 Global Hawk, and next generation experimental technologies, like the X-47B (United States) and Sharp Sword (China), demonstrate, drones are potentially useful for much more than simply targeted strikes, and in the future could engage in an even larger category of military missions. Moreover, the focus on drone strikes presumes that military robotics are only useful in the air. But there are a variety of missions – from uninhabited truck convoys to the Knifefish sea mine detection system to Israel's unmanned surface patrol vehicle, the *Protector* – in which robotic systems can play a significant role outside the context of airborne-targeted killings.<sup>5</sup>

Within the realm of military robotics, autonomy is already extensively used, including in autopilot, identifying and tracking potential targets, guidance, and weapons detonation.<sup>6</sup> Though simple autonomous

weapons are already possible, there is vast uncertainty about the state of the possible when it comes to artificial intelligence and its application to militaries. While robots that could discriminate between a person holding a rifle and a person holding a stick still seem to be on the horizon, technology is advancing quickly. How quickly and how prepared society will be for it, though, are open questions.<sup>7</sup> A small number of weapon systems currently have human-supervised autonomy. Many variants of the close-in weapon systems (CIWS) deployed by the U.S. military and more than two dozen militaries around the world, for example, have an automatic mode.<sup>8</sup> Normally, the system works by having a human operator identify and target enemy missiles or planes and fire at them. However, if the number of incoming threats is so large that a human operator cannot target and fire against them effectively, the operator can activate an automatic mode whereby the computer targets and fires against the incoming threats. There is also an override switch the human can use to stop the system.

Nearly all those discussing autonomous weapons – from international organizations to governments to the Campaign to Stop Killer Robots – agree that LAWS differ fundamentally from the weapons that militaries employ today.<sup>9</sup> While simple at first glance, this point is critical: when considering the ethical and moral challenges associated with autonomous weapons, the category only includes weapons that operate in ways appreciably *different* from the weapons of today.<sup>10</sup>

From a common sense perspective, defining an autonomous weapon as a weapon system that selects and engages targets on its own makes intuitive sense. Moreover, it is easy to describe, at the extremes, what constitutes an autonomous weapon. While a “dumb” bomb launched by a B-29 in World War II is not an autonomous weapon, a hunter-killer drone mak-

ing decisions about who to target and when to fire weapons via algorithm clearly is. In between these extremes, however, is a vast and murky gulf – from incremental advances on the precision guided weapons of today to humanoid robots stalking the earth – that complicates our thinking about the ethical and moral challenges associated with LAWS and the implications for just war theory.

In 2012, the U.S. Department of Defense (DoD) defined an autonomous weapon as “A weapon system that, once activated, can select and engage targets without further intervention by a human operator.”<sup>11</sup> The DoD further distinguished between autonomous weapons, human-supervised autonomous weapons (that is, autonomous weapons that feature a human “on the loop” who possesses an override switch), and semiautonomous weapons, or “a weapon system that, once activated, is intended to only engage individual targets or specific target groups that have been selected by a human operator.”<sup>12</sup> NGO groups, such as Human Rights Watch, have generally adopted similar definitions.<sup>13</sup> This essay does as well, considering lethal autonomous weapon systems as weapon systems that, once activated, are designed to select and engage targets not previously designated by a human.<sup>14</sup> Defining what it means to select and engage targets is complicated, however. For example, if homing munitions are considered to “select and engage” targets, then autonomous weapons have existed since World War II.

Resolving the definitional debate is beyond the scope of this essay. But even if there is not a clear agreement on exactly what constitutes an autonomous weapon, breaking down LAWS into three “types” of potential autonomous weapons – munition, platforms, and operational systems – can potentially help move the discussion forward, revealing the ethical, moral, and strategic issues that might exist for each.<sup>15</sup>

At the munitions level, there are already many semiautonomous weapons today. The advanced medium range air-to-air missile (AMRAAM), for example, deployed by the United States and several militaries around the world, is a “fire and forget” missile: after it is launched, it uses internal navigation and radar to find and destroy a target. AMRAAM engagements generally happen beyond visual range, with the pilot making the decision to launch an AMRAAM based on long-range radar data, not visual cues. The AMRAAM is not considered inherently problematic from an ethical perspective, nor is it considered an autonomous weapon.<sup>16</sup> Some fully autonomous weapons at the munitions level arguably already do exist, though, including the Israeli Harpy, a loitering cruise missile designed to detect and destroy a certain type of radar.<sup>17</sup>

The next level of military system aggregation is the platform. An example of an autonomous weapon system platform would be a ship or plane capable of selecting targets and firing munitions at those targets on its own. There are almost no platform-level LAWS currently deployed, but the CIWS systems that protect ships and military bases from attack are arguably an exception. Like the AMRAAM, countries have used these weapon systems for decades without opposition. However, an example of a platform-level LAWS that does not currently exist – and which no military appears to be planning to build – is an autonomous version of the MQ-9 Reaper (United States) or the CH-4 (China) drones. Imagine a drone identical from the exterior, but with software that allows it, after activation by a human operator, to fly around the world and target a particular individual or groups of individuals and fire missiles at them, much as human-piloted drones do today.<sup>18</sup>

The broadest type of LAWS would be a military operations planning system in which a machine learning system would

substitute, in a way, for military leaders and their staff in planning operations. No LAWS at the operational level appear to exist, even in research and development, though it is possible to imagine militaries wanting to leverage potential insights from machine learning models as they conduct planning. In this scenario, upon deciding to fight a war – or perhaps even deciding whether to fight a war – a human would activate an autonomous battle system that could decide the probability of winning a war and whether to attack, plan an operation, and then direct other systems – whether human or robotic – to engage in particular attacks. This category is the furthest away from reality in terms of technology and is the one that most invokes images of robotic weapon systems in movies such as *The Terminator* or *The Matrix*.

Some worry that autonomous weapons will be inherently difficult to use in ways that discriminate between combatants and noncombatants and only take life when necessary. An inability to discriminate would violate just war theory as well as the law of war. Consequently, some worry that autonomous weapons will be uncontrollable – prone to errors and unable to operate predictably.<sup>19</sup> Moreover, even if LAWS meet basic law of war requirements, they could create safety and control problems. Their very strength – the reliability of their programming relative to humans – could make them fragile when facing operating environments outside of their programming. At the extreme, unpredictable algorithms interacting as multiple countries deploy autonomous weapons could risk the military version of the 2010 stock market “flash crash” caused by high-frequency trading algorithms.<sup>20</sup>

Additionally, opponents of LAWS argue that autonomous weapons will necessarily struggle with judgment calls because they are not human.<sup>21</sup> For example, a human

soldier might have empathy and use judgment to decide not to kill a lawful combatant putting down a weapon or who looks like they are about to give up, while a robotic soldier would follow its order, killing the combatant. This could make it harder to use LAWS justly.<sup>22</sup>

Additionally, autonomous weapons potentially raise *jus in bello* questions concerning conduct in war from a just war perspective. For example, LAWS that are unable to respect benevolent quarantine for prisoners would violate core just war principles, though their inability to comply means responsible militaries would not deploy them in those situations. This is precisely why it makes the most sense to think about autonomous weapons in comparison with existing weapons in realistic scenarios.

These are also empirical questions, though convincing evidence is difficult to gather because these weapon systems generally do not yet exist. Moreover, even beyond the uncertainty about the technological range of the possible, many of these arguments can be made in both directions. For example, those less worried about LAWS could contend that the arguments above consider improbable scenarios, because militaries are unlikely to deploy inherently unpredictable weapons that would be less likely to accomplish missions than non-LAWS.<sup>23</sup>

In this sense, it's possible that militaries would purposefully decide not to deploy LAWS unless they believed those LAWS could operate with the ability to discriminate and follow the law of war. LAWS might also be more effective and ethical on the battlefield than other nonautonomous alternatives. Human soldiers kill unnecessarily on the battlefield, up to and including war crimes, for a variety of reasons, including rage, revenge, and errors from fatigue. One theoretical benefit of LAWS is that, as machines that do not get tired or (presumably) experience emotion, LAWS

would almost certainly fire more accurately and discriminate perfectly according to their programming. According to scholars like Ronald Arkin, this could make these types of war crimes and the killing of civilians by human soldiers less likely.<sup>24</sup>

How would these theoretical benefits and drawbacks stack up? Given the current state of the technology in question, we can only speculate the extent to which these matters are likely to be more or less serious for the three possible categories of autonomous weapon systems described above.

For munitions, most imaginable LAWS are less likely to create inherent effectiveness challenges beyond those of current weapons in terms of controllability. There is still a human operator launching the munition and making a decision about the necessity of firing upon a target or set of targets. Autonomy may help ensure that the weapon hits the correct target – or gets to the target, if autonomy enables a munition to avoid countermeasures. In this case, there is not a significant difference, from an ethical perspective, between an autonomous weapon, a semiautonomous weapon, or arguably even a bullet, because a person is making the choice to launch the munition based on what is presumably sufficient information. For example, Israel's Harpy may be problematic because the system will destroy its target whether that target is on top of a school or on a military base, but it is not executing a complicated algorithm that makes it inherently unpredictable. Practically, militaries are very unlikely to use LAWS at the munitions level unless they are demonstrably better than semiautonomous weapons, precisely for reasons of controllability.

It is, of course, possible to imagine futuristic versions of munitions that would be more complicated. Autonomous cruise missiles that can loiter for days, instead of hours, and travel around the world, programmed to target particular individuals

or ships when they meet certain criteria, could raise other questions. This is one example of how context based on geography and time may influence the appropriateness and desirability of autonomous weapon systems in a given situation.

It is at the platform and the operational levels that disquiet about discrimination and controllability becomes more complex. A LAWS platform deployed in a confined geographical space in a clear war zone may not (depending on the programming) be inherently problematic, but there are other mission sets – like patrolling autonomous drones searching for insurgents – that would lead to much greater risk from a controllability perspective. Essentially, complications, and thus the potential for fragility, will increase as the machine has to do more “work” in the area of discrimination.

At the operational battle-management level, it is difficult to imagine militaries having enough trust to delegate fundamental operational planning roles to algorithms, though they could become supplemental sources of information. Delegating those roles, however, could create large-scale ethical concerns from the consequences of those actions, in part because they might be harder to predict. Operational planning LAWS could make choices or calculate risks in novel ways, leading to actions that are logical according to their programming, but are not predictable to the humans carrying out those orders. Operational planning LAWS also connect most directly to the types of existential risks raised by Hawking and others.

One of the key arguments made by opponents of LAWS is that, because LAWS lack meaningful human control, they create a moral (and legal) accountability gap.<sup>25</sup> If they malfunction or commit war crimes, there is no single person to hold accountable the way a drone operator, pilot in the cockpit, or ground team would be account-

able today. This is potentially unique to LAWS. Remotely piloted military robotics do not appear to create excessive moral distance from war at the operator level. For example, new research shows that drone pilots actually suffer from posttraumatic stress disorder at similar rates to pilots in the cockpit.<sup>26</sup>

There is still nervousness, however, that drones already make war too “easy” for political leaders. Autonomous weapons raise similar fears, just as indirect artillery and manned airpower did in the past.<sup>27</sup> The core fear is that LAWS will allow leaders and soldiers not to feel ethically responsible for using military force because they do not understand how the machine makes decisions and they are not accountable for what the machine does.

LAWS may substitute for a human soldier, but they cannot be held accountable the way a human soldier is held accountable.<sup>28</sup> Imagine, for example, deploying a robot soldier in a counterinsurgency mission to clear a building that is suspected to house insurgents. If that robotic soldier commits a war crime, indiscriminately executing noncombatants, who is responsible? The responsible party could be the programmer, but what if the programmer never imagined that particular situation? The responsible party could be the commander who ordered the activation of the weapon, but what if the weapon behaved in a way that the commander could not have reasonably predicted?

On the other side of the debate, part of the problem is imagining LAWS as agents, rather than tools. The human operator that fires a LAWS munition or activates a LAWS platform still has an obligation to ensure the system will perform in an ethically appropriate fashion to the best of anyone’s ability to predict, just as with today’s weapons.<sup>29</sup> Thus, planning and training becomes critical to avoiding a responsibility gap. By ensuring that potential op-

erators of LAWS understand how they operate – and feel personally accountable for their use – militaries can theoretically avoid offloading moral responsibility for the use of force.

Formal rules could ensure technical accountability. One solution in the case of the ground combat situation described above is to hold the commander accountable for war crimes committed by the robotic soldier, just as commanders today are generally held accountable for war crimes committed by their unit.<sup>30</sup> This leads to fairness considerations, though: if the robotic soldier malfunctions, and it is not the fault of the commander, is it fair to hold the commander accountable? Arguably not, though commander accountability for LAWS would create a strong incentive for commanders only to use LAWS when they have a high degree of confidence in its situational appropriateness. Analogies from legal regimes, such as vicarious liability, could also prove useful. Thus, while accountability and responsibility issues are relevant topics, it is not clear that they are irresolvable. Additionally, accidents with nonautonomous and semiautonomous weapons happen today and raise accountability questions. In a 2003 incident in which a U.S. Patriot missile battery shot down allied aircraft, no one was personally held accountable for the system malfunction. Should the accountability requirements for LAWS be higher than for other weapon systems?

Considering this argument in both directions, it makes sense again to see how these concerns might vary across different types of LAWS. At the munitions level, ensuring legal accountability and moral responsibility should be relatively close, if not identical, to the use of semiautonomous weapons today. There will still be human operators firing the munitions in ways that they believe are legitimate; the guidance systems for the munitions would

just operate somewhat differently. Adaptations of existing accountability regimes therefore seem plausible.

The platform level will place the largest amount of stress on potential training and planning to avoid offloading accountability when using LAWS. While there is still a person that will have to activate and launch an autonomous weapons platform, if that person lacks sufficient understanding of the mission or how the LAWS will operate to complete the mission, it could lead to a responsibility gap. Such a gap does not seem inevitable, however, presuming the construction of clear rules and training.

At the operational system level, the use of LAWS creates a real and significant risk of moral offloading. Operational planning conducted by an algorithm – rather than the algorithm being one input into human judgment – is precisely the type of situation in which human accountability for war would decline and humans might cease to feel responsible for the casualties caused by war. This is a significant ethical concern on its own and would raise large questions in terms of just war theory.

Establishing the line at which the human is so removed from the targeting decision that it makes the use of force *a priori* unjust is complex from a just war perspective, however. Imagine a case in which the human is entirely removed from the targeting and firing process, but the outcome is a more precise military engagement. On the one hand, such an engagement would almost certainly meet basic *jus in bello* requirements, but one might also argue that the removal of human agency from the process is ethically defective. This is a tricky question, and one worth further consideration.

The last major ethical argument about LAWS is whether they might be inherently problematic because they dehumanize their targets. All human life is precious and has intrinsic value, so having machines

select and engage targets arguably violates fundamental human dignity – people have the right to be killed by someone who made the choice to kill them. Since machines are not moral actors, automating the process of killing through LAWS is also by definition unethical, or as technology philosopher Peter Asaro has put it: “justice itself cannot be delegated to automated processes.”<sup>31</sup> LAWS might therefore be thought of as *mala in se*, or evil in themselves, under just war theory.

If a machine without intentions or morality makes the decision to kill, it makes us question why the victim died.<sup>32</sup> This argument has a moral force. As human rights legal scholar Christof Heyns argues: “Decisions over life and death in armed conflict may require compassion and intuition.”<sup>33</sup> There is something unnerving about the idea of machines making the decision to kill. The United Nations Institute for Disarmament Research describes it as “an instinctual revulsion against the idea of machines ‘deciding’ to kill humans.”<sup>34</sup> The concern by opponents of LAWS is that machines making decisions about killing leads to a “vacuum of moral responsibility”: the military necessity of killing someone is a subjective decision that should inherently be made by humans.<sup>35</sup>

On the other side, all who enter the military understand the risks involved, including the potential to die; what difference does the *how* make once you are dead? In an esoteric sense, there may be something undignified about dying at the hands of a machine, but why is being shot through the head or heart and instantly killed by a machine necessarily worse than being bludgeoned by a person, lit on fire, or killed by a cruise missile strike? The dignity argument has emotional resonance, but it may romanticize warfare. Humans have engaged in war on an impersonal and industrial scale since at least the nineteenth century: from the near sixty thousand

British casualties the first day of the Battle of the Somme to the firebombing of Tokyo and beyond.

Looking at the three categories of possible LAWS again reveals potential differences between them with regards to the question of human dignity. At the munitions level, LAWS seem unlikely to generate significant human dignity questions beyond those posed by existing weapon systems, at least based on the current technological world of the possible. Since the decision-making process for the use of force would be similar, if not identical, to the use of force today, the connection between the individual firing the weapon and those affected would not change.<sup>36</sup>

At the platform level, LAWS again require deeper consideration, because it is with LAWS platforms that the system begins calculating whether to use force. The extent to which they may be problematic from a human dignity perspective may also again depend on how they are used. The use of platform-level LAWS in an anti-material role against adversary ships or planes on a clear battlefield would be different than in an urban environment. Moreover, as the sophistication of LAWS grows, they could increase the risk of dehumanizing targets. Returning to the case of the Harpy, at present, it is clearly up to the person launching the missile to make sure there is a lawful radar target that the Harpy can engage. A platform with the ability to make choices about whether the radar is a lawful target (for example, is the radar on top of a hospital?) would be better at discrimination, making it ethically preferable in some ways, but also raising questions from the perspective of the human dignity argument; it is the machine, rather than a person, making the targeting decision.<sup>37</sup>

The human dignity argument arguably also applies less to platforms that defend a fixed position from attack. Electric fences are not ethically problematic as a cat-



egory if labeled clearly and used in areas where any intrusion is almost by definition a hostile action.<sup>38</sup> Or to take another example, South Korea deploys a gun system called the SGR-1 pointed at the demilitarized zone with North Korea. The system has some automatic targeting features, though the specifics are unclear. However, since the system is deployed in a conflict zone and can only aim at targets that would almost certainly be lawful combatants, this is arguably less problematic than LAWS platforms employed as part of an assault operation.

LAWS pose the largest challenges to human dignity at the operational system level, though the relationship to just war theory is more ambiguous. An operational-level LAWS making decisions about whether and how to conduct a military operation certainly involves offloading moral responsibility for the use of force to a machine. Oddly, though, imagine a case in which an operational-level LAWS designed a battle plan implemented by humans. In that case, the machine is taking the place of a high-level military commander, but humans are selecting and engaging targets on the ground. Would this be less problematic, ethically, than a hunter-killer drone searching for individuals or groups of insurgents? It sounds odd, but this example points to the complexities of assessing these issues.

The debate is just beginning, and this essay attempts to address the broad ethical issues potentially associated with the development of autonomous weapons, a class of weapons that, with a few exceptions, do not yet exist. While technological trends suggest that artificial intelligence is rapidly advancing, we are far from the realm of dystopian science fiction scenarios. Of course, how quickly the technology will develop remains to be seen.

Do autonomous weapons create novel issues from an ethical perspective, espe-

cially regarding just war theory? Excluding technologically implausible scenarios of autonomous operational battle systems deciding to go to war, autonomous weapons are unlikely to lead to *jus ad bellum* problems from a traditional just war perspective, excluding the risk that LAWS will make going to war so easy that political leaders will view unjust wars as costless and desirable. One could argue that since machines cannot have intentions, they cannot satisfy the *jus ad bellum* requirement for right intentions. Yet this interpretation would also mean that broad swaths of precision-guided modern semiautonomous weapons that dramatically reduce civilian suffering in war arguably violate the individual intentionality proposition, given the use of computerized targeting and guidance. Presumably no one would rather the world return to the age of the “dumb bombs” used in World War II. Overall, it is critical to understand that there is the possibility for significant diversity within the subset of autonomous weapons, in particular, whether one is discussing a munition with greater autonomy in engaging a target versus a platform or operational system.

At the level of the munition, where LAWS might represent missiles programmed to attack particular classes of targets (such as an amphibious landing craft) in a given geographic space, the relevant ethical issues appear similar to those regarding today’s weapons. The process of using force – and responsibility for using force – would likely look much the same as it does today for drone strikes or the use of other platforms that launch precision-guided munitions. The key will be how munitions-based LAWS are used.

It is at the platform level that the ethical challenges of LAWS begin to come into focus. Autonomous planes, for example, flying for thousands of miles and deciding for themselves whom to target, could risk the moral offloading of responsibility and un-

dermine human dignity in some scenarios, even if they behave in ways that comply with the law of war. While it is possible to address this issue through training, accountability rules, and restricting the scenarios for using autonomous weapon platforms, this area requires further investigation.

Autonomous operational systems using algorithms to decide whether to fight and how to conduct operations, besides being closest to the robotic weapon systems of movies and television, could create more significant moral quandaries. Given full authority (as opposed to supplementing human judgment), operational system LAWS would make humans less relevant, from an ethical perspective, in major wartime decision-making. Fortunately, these types of systems are far from the technological range of the possible, and humans

are quite unlikely to want to relinquish that level of control over war, meaning the real world systems that require deeper thought over the next several years are LAWS at the munition and platform levels.

Finally, just war theory provides an interesting lens through which to view LAWS: could they lead to a world in which humans are more removed from the process of warfare than ever before, while warfare itself becomes more precise and involves less unnecessary suffering? These are complicated questions regarding the appropriate role for humans in war, informed by how we balance evaluating LAWS based on a logic of consequences versus evaluating them based on a logic of morality. It will be critical to ensure in any case that the human element remains a central part of warfare.

#### ENDNOTES

Author's Note: Thank you to Michael Simon and all the workshop participants at West Point, along with Paul Scharre, for their feedback. All errors are the sole responsibility of the author.

- <sup>1</sup> Michael C. Horowitz, Sarah E. Kreps, and Matthew Fuhrmann, "The Consequences of Drone Proliferation: Separating Fact from Fiction," working paper (Philadelphia: University of Pennsylvania, 2016).
- <sup>2</sup> For the purposes of this essay, I use the phrases *autonomous weapon*, *autonomous weapon system*, and *lethal autonomous weapon system* interchangeably.
- <sup>3</sup> See the Campaign to Stop Killer Robots, <http://www.stopkillerrobots.org/>; and the Future of Life Institute, "Autonomous Weapons: An Open Letter from AI and Robotics Researchers," [http://futureoflife.org/AI/open\\_letter\\_autonomous\\_weapons](http://futureoflife.org/AI/open_letter_autonomous_weapons).
- <sup>4</sup> For example, see the discussion in Peter Asaro, "On Banning Autonomous Weapon Systems: Human Rights, Automation, and the Dehumanization of Lethal Decision-Making," *International Review of the Red Cross* 94 (886) (2012): 687–709; and Charli Carpenter, "How Do Americans Feel About Fully Autonomous Weapons?" *Duck of Minerva*, June 10, 2013, <http://duckofminerva.com/2013/06/how-do-americans-feel-about-fully-autonomous-weapons.html>; and Michael N. Schmitt, "Autonomous Weapon Systems and International Humanitarian Law: A Reply to the Critics," *Harvard National Security Journal* (2013); and Michael C. Horowitz, "Public Opinion and the Politics of the Killer Robots Debate," *Research & Politics* (forthcoming).
- <sup>5</sup> Michael C. Horowitz, "The Looming Robotics Gap: Why America's Global Dominance in Military Technology is Starting to Crumble," *Foreign Policy Magazine* (May/June 2014), <http://foreignpolicy.com/2014/2005/2005/the-looming-robotics-gap/>. Put another way, discussions of banning drones because they are used for targeted killing conflate the act of concern (targeted killings) with the means (drones), when other means exist. It would be like banning the airplane in the early twentieth century because of targeted killing.

- <sup>6</sup> Paul Scharre and Michael C. Horowitz, “An Introduction to Autonomy in Weapon Systems,” *Michael C. Horowitz* working paper (Washington, D.C.: Center for a New American Security, 2015), 7, [http://www.cnas.org/sites/default/files/publications-pdf/Ethical%20Autonomy%20Working%20Paper\\_021015\\_v02.pdf](http://www.cnas.org/sites/default/files/publications-pdf/Ethical%20Autonomy%20Working%20Paper_021015_v02.pdf).
- <sup>7</sup> For one example, see Stuart Russell, “Artificial Intelligence: Implications for Autonomous Weapons,” presentation at the Convention on Certain Conventional Weapons, Geneva, 2015.
- <sup>8</sup> U.S. military examples include the Phalanx and C-RAM.
- <sup>9</sup> Human Rights Watch, *Losing Humanity: The Case against Killer Robots* (New York: Human Rights Watch, 2012), [http://www.hrw.org/sites/default/files/reports/arms1112\\_ForUpload.pdf](http://www.hrw.org/sites/default/files/reports/arms1112_ForUpload.pdf).
- <sup>10</sup> It is possible, of course, to use today’s weapons in ethically problematic ways, but that is beyond the scope of this essay.
- <sup>11</sup> U.S. Department of Defense, *Directive on Autonomy in Weapons Systems, Number 3000.09* (Washington, D.C.: U.S. Department of Defense, 2012), 13, <http://www.dtic.mil/whs/directives/corres/pdf/300009p.pdf>.
- <sup>12</sup> *Ibid.*, 14.
- <sup>13</sup> Human Rights Watch, *Losing Humanity*.
- <sup>14</sup> This builds on the definition in Scharre and Horowitz, “An Introduction to Autonomy in Weapon Systems,” 16. The phrase “not previously designated by a human” helps reconcile the fact that the use of weapons sometimes involves firing multiple munitions at multiple targets.
- <sup>15</sup> Another interesting possibility is to classify LAWS based on the types of autonomy they possess. See Heather Roff, “The Forest for the Trees: Autonomous Weapons and ‘Autonomy’ in Weapons Systems,” working paper, June 2016.
- <sup>16</sup> This discussion is similar to *ibid.*, 11.
- <sup>17</sup> Peter J. Spielmann, “Israel Killer Robots Could be Banned under UN Proposal,” *The Times of Israel*, May 3, 2013, <http://www.timesofisrael.com/israeli-killer-robots-could-be-banned-under-un-proposal/>.
- <sup>18</sup> The X-47B, a U.S. Navy experimental drone, has autonomous piloting, but not automated weapon systems.
- <sup>19</sup> Human Rights Watch, *Losing Humanity*.
- <sup>20</sup> Michael C. Horowitz and Paul Scharre, “The Morality of Robotic War,” *The New York Times*, May 26, 2015, <http://www.nytimes.com/2015/05/27/opinion/the-morality-of-robotic-war.html>. Also see Paul Scharre, *Autonomous Weapons and Operational Risk* (Washington, D.C.: Center for a New American Security, February 2016), [http://www.cnas.org/sites/default/files/publications-pdf/CNAS\\_Autonomous-weapons-operational-risk.pdf](http://www.cnas.org/sites/default/files/publications-pdf/CNAS_Autonomous-weapons-operational-risk.pdf).
- <sup>21</sup> Aaron M. Johnson and Sidney Axinn, “The Morality of Autonomous Robots,” *Journal of Military Ethics* 12 (2) (2013): 137.
- <sup>22</sup> Michael Walzer, *Just and Unjust Wars: A Moral Argument with Historical Illustrations*, 4th ed. (New York: Basic Books, 1977), 142 – 143.
- <sup>23</sup> This is particularly true given that drones and other remotely piloted military robotics options exist.
- <sup>24</sup> Ronald C. Arkin, *Governing Lethal Behavior in Autonomous Robots* (Boca Raton, Fla.: CRC Press, 2009).
- <sup>25</sup> Wendell Wallach, *A Dangerous Master: How to Keep Technology from Slipping Beyond Our Control* (New York: Basic Books, 2015); and Human Rights Watch, *Mind the Gap: The Lack of Accountability for Killer Robots* (New York: Human Rights Watch, 2015), <https://www.hrw.org/report/2015/04/09/mind-gap/lack-accountability-killer-robots>.

- <sup>26</sup> James Dao, “Drone Pilots are Found to Get Stress Disorders Much as Those in Combat Do,” *The New York Times*, February, 22 2013, <http://www.nytimes.com/2013/02/23/us/drone-pilots-found-to-get-stress-disorders-much-as-those-in-combat-do.html>.
- <sup>27</sup> Kenneth Anderson, Daniel Reisner, and Matthew C. Waxman, “Adapting the Law of Armed Conflict to Autonomous Weapon Systems,” *International Law Studies* 90 (2014): 391 – 393; and Kenneth Anderson and Matthew C. Waxman, “Law and Ethics for Autonomous Weapon Systems: Why a Ban Won’t Work and How the Laws of War Can,” *Jean Perkins Task Force on National Security and Law Essay Series* (Stanford, Calif.: Stanford University, The Hoover Institution, April 10, 2013).
- <sup>28</sup> Robert Sparrow, “Killer Robots,” *Journal of Applied Philosophy* 24 (1) (2007): 62 – 77.
- <sup>29</sup> Horowitz and Scharre, “The Morality of Robotic War.”
- <sup>30</sup> This can vary depending on the specific situation, but the general point is clear.
- <sup>31</sup> Asaro, “On Banning Autonomous Weapon Systems,” 701.
- <sup>32</sup> United Nations Institute for Disarmament Research, *The Weaponization of Increasingly Autonomous Technologies: Considering Ethics and Social Values* (Geneva: United Nations Institute for Disarmament Research, 2015), 9, <http://www.unidir.org/files/publications/pdfs/considering-ethics-and-social-values-en-624.pdf>.
- <sup>33</sup> United Nations General Assembly, *Report of the Special Rapporteur on Extrajudicial, Summary or Arbitrary Executions (A/HRC/23/47)*, April 9, 2013, [http://www.ohchr.org/Documents/HRBodies/HRCouncil/RegularSession/Session23/A-HRC-23-47\\_en.pdf](http://www.ohchr.org/Documents/HRBodies/HRCouncil/RegularSession/Session23/A-HRC-23-47_en.pdf).
- <sup>34</sup> United Nations Institute for Disarmament Research, *Weaponization of Increasingly Autonomous Technologies*, 7 – 8.
- <sup>35</sup> United Nations General Assembly, *Report of the Special Rapporteur*, 17.
- <sup>36</sup> This is arguably why munitions-based LAWS may not really be LAWS at all, depending on the definition.
- <sup>37</sup> Thanks to Paul Scharre for making this point clear to me in a personal conversation.
- <sup>38</sup> Johnson and Axinn, “The Morality of Autonomous Robots,” 131.