

The Ethics of Acquiring Disruptive Technologies

Artificial Intelligence, Autonomous Weapons, and Decision Support Systems

By C. Anthony Pfaff

Last spring, Google announced that it would not partner with the Department of Defense (DOD) on “Project Maven,” which sought to harness the power of artificial intelligence (AI) to improve intelligence collection and targeting, because its employees did not want to be “evil.”¹ Later that fall, the European Union called for a complete ban on autonomous weapons systems.² In fact, a number of AI-related organizations and researchers have signed a “Lethal Autonomous Weapons Pledge” that expressly prohibits development of machines that can decide to take a human life.³

Reluctance to develop AI applications for military purposes is not going to go away as the development, acquisition, and employment of these systems challenge the traditional norms associated with not just war-fighting but morality in general.⁴ Meanwhile, as the debate rages, adversaries of the United States who do not have these ethical concerns continue with their development. China, for example, has vowed to be the leader in AI by 2030.⁵ No one should have any illusions that the Chinese will not use this dominance for military as well as civilian purposes. So, to maintain parity, if not advantage, DOD has little choice but to proceed with the development and employment of artificially intelligent systems. As it does so, ethical concerns will continue to arise, potentially excluding important expertise for their development. To include this expertise, DOD needs to confront these concerns upfront.

Because this technology challenges traditional norms, it is disruptive in ways that more conventional technologies are not. This disruption arises because these technologies do not simply replace older ones but change how actors compete.⁶ Changing how actors compete in effect changes the game—which, in turn, changes the rules. To effectively compete in the new environment, actors then have to establish new rules. For example, the development of the internet changed how people obtained news and information, forcing the closure of more traditional media, which struggled to find ways to generate revenue under these new conditions.

Of course, in addressing ethical concerns, it is important to be clear regarding the source of the “evil” in question. To the extent opponents of the technology are committed pacifists, then there is little reason one could bring to bear to change their minds. Another way, however, of understanding their objection is not that all military research is evil, but that development and use of AI weapons systems are *mala en se*, which means

Dr. C. Anthony Pfaff is a Research Professor for Strategy, the Military Profession, and Ethics, Strategic Studies Institute, U.S. Army War College.

their use, in any context, constitutes a moral wrong. If true, such weapons would fall into the same category as chemical and biological weapons, the use of which is banned by international law. If these weapons do fit into that category, the U.S. Government's only morally appropriate response would be to work to establish an international ban rather than to develop them.

The difficulty here is that no one really knows if these weapons are inherently evil. Objections to their use tend to cluster around the themes that such weapons introduce a "responsibility gap" that could undermine international humanitarian law (IHL) and dehumanize warfare in ways that are morally unacceptable. Moreover, even if one resolved these concerns, the application of such systems risks moral hazards associated with lowering the threshold to war, desensitizing soldiers to violence, and incentivizing a misguided trust in the machine that abdicates human responsibility. At the same time, however, proponents of such systems correctly point out that they are not only typically more precise than their human counterparts, they also do not suffer from emotions such as anger, revenge, and frustration that give rise to war crimes.

The answer, of course, will depend on how DOD proceeds with the development of these systems. While the responsibility gap and dehumanization of warfare are important moral objections, there are ways to address them as well as the moral hazards to which these systems give rise. Addressing these objections, however, will require attention throughout the development, acquisition, and employment cycles.

For the purposes of this discussion, the term *AI systems* will refer to military AI systems that may be involved in life-and-death decisions. These systems include both lethal autonomous weapons systems (LAWS) that can select and engage targets without human intervention, and decision support systems (DSS) that facilitate complex decisionmaking

processes, such as operational and logistics planning. After a brief discussion of military applications of AI, I will take on the question of whether these systems are *mala en se* and argue that the objections described above are insufficient to establish that they are inherently evil. Still, this is new and disruptive technology that gives rise to moral hazards that are unique to its employment. I will address that concern and discuss measures DOD can take to mitigate these hazards so that the employment of these systems conforms to our moral commitments.

The Responsibility Gap

Advocates of AI systems often make the point that these systems' capabilities enable ethical behavior better than that of human beings, especially in combat. Ronald G. Arkin, in his book *Governing Lethal Behavior in Autonomous Robots*, argues that not only do such systems often have greater situational awareness, they also are not motivated by self-preservation, fear, anger, revenge, or misplaced loyalty, suggesting that they will be less likely to violate the war convention than their human counterparts.⁷ Improving ethical outcomes, however, is only part of the problem. While machines may perform ethically better than humans—in certain conditions at least—they still make mistakes.

The fact that machines can make mistakes entails the possibility for harm to persons for which no one is accountable. This point, however, is not intuitively obvious. Conventional weapons can also malfunction and result in unjust death or harm. That fact, however, does not generate concern regarding the rules for their employment or raise ethical concerns regarding their acquisition. As operators can trust these systems to generally function reliably and their effects conform to IHL, then usually there are no ethical reasons—beyond a general commitment to pacifism—to oppose their employment. When the employment of such systems does result in excessive or unjustified harm,

it is possible conceptually, if not always practically, to determine where the responsibility lies: with the operator, who may have used the system improperly; the manufacturer, which may have made an error in the systems construction; or the developer, who may have designed the system poorly. The source, whatever it is, is human, and humans can be held accountable for their actions.

AI-driven systems, however, erode that trust in two ways. First, rather than simply being instrumental to human decisions, they take some of the burden of that decisionmaking off of humans. Moreover, as the technology becomes more effective, humans will likely become more dependent on these machines for making life-and-death decisions. In Iraq, for example, some units employed a DSS to select the safest route for convoys to travel. It made this recommendation based on attack and other incident data it received.

In one instance, the machine recommended a route on which a convoy suffered an attack where U.S. soldiers' lives were lost. As it turned out, the recommended route had previously been categorized as one of the most dangerous routes in the area of operations. Because it was one of the more dangerous routes, convoys stopped using it; thus, over time, it appeared from the perspective of the DSS as one of the safer routes since there had not been any recent recorded attacks.⁸ This is, of course, a simple example from a very rudimentary DSS. However, it does raise the question: on what basis can an operator reasonably trust an AI-driven system?

Second, as these machines become more complex, it may not always be possible to tell why they behaved the way they did. For example, in June 2007, the first three "Talon" Special Weapons Observation Reconnaissance Detection Systems—remotely controlled robots that can aim at targets automatically—were deployed to Iraq but reportedly never used because their guns moved when they should not have.⁹

To mitigate the hazards of unaccountable moral failure, one has to ensure that either a human makes life-and-death decisions or that machines develop the capability to act as "autonomous moral agents." Both give rise to moral concerns unique to AI systems. Keeping a human in the decisionmaking process often means that one cannot take full advantage of the capabilities of LAWS and DSS.¹⁰ When that decision is to fire at a rapidly approaching missile, any delay can make the difference between life and death. Additionally, employing such systems risks desensitizing soldiers, lowering the threshold for violence, and developing a bias favoring machine judgments over human ones, which I will take up later.

To understand the difficulty in resolving these concerns, it will help to understand what it would take to turn AI systems into what Wendell Wallach and Colin Allen call "autonomous moral agents." They argue in their book, *Moral Machines: Teaching Robots Right from Wrong*, that moral agents require the ability to "monitor and regulate their behavior in light of the harms their actions may cause or the duties they may neglect."¹¹ Moreover, they further require the ability to "detect the possibility of harm or neglect of duty" as well as to take steps to minimize or avoid any undesirable outcomes.¹² There are two routes to achieving this level of agency. First is for designers and programmers to anticipate all possible courses of action and determine rules that result in desired outcomes in the situations in which the autonomous system will be employed. Second is to build "learning" systems that can gather information, attempt to predict consequences of their actions based on that information, and determine an appropriate response.¹³ The former requires either a great deal of knowledge on the part of the programmer or a very limited application for the machine. The latter requires overcoming the related problems of ascription and isotropy.

Ascription refers to the way humans infer other persons' intent from their actions. *Isotropy* refers to

the human ability to determine the environmental elements or background knowledge relevant to that ascription. Consider, for example, a group of soldiers who burst into a room looking for an insurgent and see several young children with knives running in their direction. On what basis do they consider the children a threat? If the operation were conducted in a Sikh village, these soldiers might know that Sikhs often wear a traditional dagger, which is symbolic and never used as a weapon. Moreover, they might be able to discern from the way the children were running and other relevant environmental cues that the children were at play and not to be taken as a threat.¹⁴

To program that capability in a robot would also be a daunting, perhaps impossible, challenge. Ascribing mental states to others requires humans to see in others the beliefs, desires, hopes, fears, intentions, and so on, that they see in themselves. As Marcello Guarini and Paul Bello observe, the relevant information associated with such ascriptions is extensive and includes such diverse things as “facial expressions, gaze orientation, body language, attire, information about the agent’s movement through an environment, information about the agent’s sensory apparatus, information about the agent’s background beliefs, desires, hopes, fears, and other mental states.”¹⁵

This difficulty is frequently referred to as the frame problem. Human knowledge about the world is holistic, where changes in one bit of knowledge can affect others. For example, learning that one is out of milk may require one to schedule a grocery shopping trip, which in turn might cause one to reschedule a meeting as well as determine to buy more milk than one had previously done. While humans do this easily, a computational AI system must go through all of its stored information to test how running out of milk affects it.¹⁶

Even when an AI system can reasonably handle sorting through alternatives, its output—whether

it is behavior or a course of action—lacks intention, an important component to moral analysis. In his famous “Chinese room” thought experiment, philosopher John Searle described a man who sits in a room. His job is to take input, in the form of Chinese characters, and consult a rule book that tells him what the output should be. He then takes the appropriate characters and provides them to whom-ever is outside the room. He does not understand the meaning of the characters, only how they relate to the rules. Thus, given a sufficiently complex rule book, he conceptually can mimic a fluent Chinese speaker without understanding anything being said.¹⁷ Moreover, while he may be causally responsible for the output, since he does not understand it, he cannot be said to intend its content. If he does not intend the content of the response, it can further be concluded that he is indifferent to it.

The point here is that one thing at least that differentiates humans from machines is that humans, in the words of AI theorist John Haugeland, “give a damn.” He argues that understanding language depends on caring about not only one’s self, but also the world in which one lives.¹⁸ The human in the Chinese room may care (or at least has the ability to care) that he gets the rules right, but the machine itself does not have the capacity to care what the output actually means. There is a difference between something being manipulated according to a set of rules and someone acting on one’s volition according to rules.¹⁹ The output of the Chinese room is clearly an example of the former.

This point suggests that AI systems will be limited in how they can interpret a complex environment, inviting error in their decisionmaking even when all the humans involved in their employment have done everything right. As Wallach and Allen also observe, “As either the environment becomes more complex or the internal processing of the computational system requires the management of a wide array of variables, the designers and

engineers who built the system may no longer be able to predict the many circumstances the system will encounter or the manner in which it will process new information.”²⁰ If it is not possible to fully account for machine behavior in terms of decisions by human beings, then it is possible to have an ethical violation for which no one is responsible.

Further complicating accountability is the fact that not all AI-driven behavior is attributable to written code. As Paul Scharre and Michael Horowitz observe, at least some of the information certain AI systems use to determine their responses is often encoded in the strength of connections of their neural networks and not as code that human operators can analyze. As a result, machine thinking can be something of a “black box.”²¹ Thus, the inability to fully account for machine behavior introduces a “responsibility gap,” which threatens to undermine the application of the war convention.²² This gap is not just a function of accessibility and complexity. It also follows from the fact these machines can make mistakes or commit an unjustified harm even when they are functioning properly. Thus, the developer, manufacturer, and operator can do all the right things and unjustified harm can still be done.

To the extent AI systems cannot be morally responsible for the harm they cause, Hin-Yan Liu views the application of these systems as the moral equivalent of employing child soldiers who are also not morally responsible for the harm they commit. Rather, he argues, since international law criminalizes the introduction of children on the battlefield, regardless of how they behave, the crime for those who employ them is not simply that they are victimizing children, but that they are also introducing “irresponsible entities” on the battlefield. Since AI systems are also “irresponsible” in the relevant sense, they too should be banned and those who do introduce them subject to criminal penalty.²³

The concern here is that accountability is a critical aspect of any normative regime. Norms,

whatever form they come in, moral, legal, or practical, are the means by which we communicate to others that we hold them, and ourselves, accountable. However, when norms are not upheld, they die.²⁴ Consider a workplace environment where there is a norm, for example, to show up on time. If workers instead habitually show up late and are not held accountable, then they will likely continue to do so and others are likely to follow. Eventually, the norm to show up on time will cease to be a norm.

The employment of non-accountable AI systems risks the same fate to IHL as well as any other regulatory scheme governing the military. The fact that LAWS and DSS can absolve humans of accountability for at least some violations will establish an incentive to employ the machines more often and find ways to blame them when something goes wrong, even when a human is actually responsible. It is not hard to imagine that over time, there would be sufficient unaccountable violations that the rules themselves would rarely be applied, even to humans. This point suggests that it will be insufficient to defend the use of AI systems simply because they can be necessary, proportionate, discriminate, and avoid unnecessary suffering if their use threatens to undermine the rules themselves.²⁵

While Liu is right that the utilization of autonomous systems does introduce entities that cannot be held accountable for harms they commit, it is worth considering the significance of the objection. One response could be to note that while with conventional weapons humans are typically responsible for any harms, it is not the case that they are always *held* responsible. That being the case, it is not clear there is a moral difference between failing to hold someone responsible and there not being someone to hold responsible. It would seem both conditions are equally destructive to a normative framework.

The difference, of course, lies in the options one has to remedy the situation. The appropriate response to failing to hold wrongdoers accountable

is, obviously, to overcome whatever barriers there are to accountability. However, in cases where there is no one to be held accountable, the way forward is not so clear.

Whatever that way forward is, the first step would seem to lie in finding a way to establish “meaningful human control” over these systems. Unfortunately, there is no set standard for meaningful human control that could apply. At one extreme, activist groups such as the International Committee for Robot Arms Control argue that meaningful human control must entail human operators having full contextual and situational awareness of the target area. They also need sufficient time for deliberation on the nature of the target, the necessity and appropriateness of attack, and the likely collateral harms and effects. Finally, they must have the means to abort the attack if necessary to meet the other conditions.²⁶

The difficulty with these criteria is that they hold the systems to a higher standard than non-autonomous weapons systems already in use. Soldiers and their commanders rarely have “full contextual and situational awareness of a target area.” Even when they do, soldiers who fire their rifles at an enemy have no ability to prevent the bullet from striking wherever they aimed it. It seems odd, then, to ban future weapons based on higher standards than the ones that current weapons meet.²⁷ It makes less sense when one realizes that some of the capabilities that come along with autonomous weapons can set conditions for better moral decisionmaking.

So whatever standard for meaningful human control one employs, it should reflect the abilities as well as limitations humans actually have. To the extent the problem lies in the machine’s ability to displace or undermine human agency, the remedy lies in restoring it. For AI-driven systems, this remedy entails diffusing responsibility throughout the development, production, and employment processes. To do so, DOD should ensure the following:

- acquisition officials, designers, programmers, and manufacturers, as well as commanders and operators, must fulfill their roles with the war convention in mind;
- commanders and operators must be knowledgeable not only regarding what the machine is doing, they also must be sufficiently knowledgeable regarding how the machine works so they better understand how it will interpret and act on instructions as well as provide output;
- commanders and operators must be in a position to prevent machine violations, either by ensuring they authorize all potentially harmful actions by the machine or by being able to monitor the operations of the machine and prevent them from happening; and
- systems in which operator intervention is not possible should only be employed in situations where commanders and operators can trust them to perform at least as well as human soldiers.

Such measures may represent the best humans can do to limit mistakes, but they will not *eliminate* them. Given that the possibility for unaccountable error endures, should we then declare AI-driven systems *mala en se*? The short answer is no. As Geoffrey S. Corn argues, while humanitarian constraints on the conduct of war are a “noble goal,” they do not exhaust the war convention, which permits states to defend themselves and others from aggression. As Corn notes;

When these constraints are perceived as prohibiting operationally or tactically logical methods or means of warfare, it creates risk that the profession of arms—the very constituents who must embrace the law—will see it as a fiction at best or, at worst, that will feign commitment to the law while pursuing sub rosa agendas to sidestep obligations.²⁸

The concern here is not that soldiers will side-step legal or moral obligations because upholding them represents excessive risk. There are, however, deeper obligations at play. States have an obligation not only to defend their citizens but also to ensure that those citizens who come to that defense have every advantage to do so successfully at the least risk possible.²⁹ Thus, any ban on LAWS or DSS, to the extent it limits chances for success or puts soldiers at greater risk, represents its own kind of moral failure.

While this point suggests it is premature to declare AI-driven systems inherently evil, it is also insufficient to fully establish their permissibility. It is not enough to point out that AI systems can lead to better ethical outcomes without fully accounting for the unethical ones. It may be permissible to accept some ethical “risk” regarding human incentives as these can be compensated for by additional rules and oversight. When those are inadequate, as I will discuss later, there are still other ways to address the responsibility gap given any particular human-machine relationship. This point suggests that where humans can establish sufficient control over these systems to be responsible for their behavior, their use would be permissible. The difficulty with this approach, however, is that such control usually comes at the expense of using this technology to its full capability. Moreover, such control over the machines ignores the impact on humans.

Dehumanizing Warfare

On the surface, concerns about dehumanizing warfare seem odd. War may be a human activity, but rarely does it feel to those involved like a particularly humane activity, often bringing out the worst in humans rather than the best. Thus, if LAWS and DSS can reduce some of the cruelty and pain war inevitably brings, it is reasonable to question whether dehumanizing war is really a bad thing. As Paul Scharre notes, the complaint that respecting human dignity requires that only humans

make decisions about killing “is an unusual, almost bizarre critique of autonomous weapons . . . there is no legal, ethical, or historical tradition of combatants affording their enemies the right to die a dignified death in war.”³⁰

Scharre’s response, however, misses the point. He is correct that AI systems do not represent a fundamentally different way for enemy soldiers and civilians to die than those that human soldiers are permitted to employ. The concern here, however, is not that death by robot represents a more horrible outcome than when a human pulls the trigger. Rather, it has to do with the nature of morality itself and the central role that respect for persons, understood in the Kantian sense as something moral agents owe each other, plays in forming our moral judgments.

Drawing on Kant, Robert Sparrow argues that respect for persons entails that even in war, one must acknowledge the personhood of those one interacts with, including the enemy. Acknowledging that personhood requires that whatever one does to another, it is done intentionally, with the knowledge that whatever the act is, it is affecting another person.³¹ This relationship does not require communication or even the awareness by one actor that he or she may be acted upon by another. It just requires that the reasons actors give for any act that affects another human being take into account the respect owed that particular human being. To make life-and-death decisions absent that relationship subjects human beings to an impersonal and predetermined process, and subjecting human beings to such a process is *disrespectful* of their status as human beings.

Thus, a concern arises when *non-moral* agents impose *moral* consequences on moral agents. Consider, for example, an artificially intelligent system that provides legal judgments on human violators. It is certainly conceivable that engineers could design a machine that could take into account a larger quantity and variety of data than could a human judge. The difficulty with the judgment the machine

renders, however, is that the machine cannot put itself in the position of the person it is judging and ask, “If I were in that person’s circumstances, would I have done the same thing?” It is the inability to not only empathize but then employ that empathy to generate additional reasons to act (or not act) that makes the machine’s judgment impersonal and predetermined.³²

Absent an interpersonal relationship between judge and defendant, defendants have little ability to appeal to the range of sensibilities human judges may have to get beyond the letter of the law and decide in their favor. In fact, the European Union has enshrined the right of persons not to be subject to decisions based solely on automated data processing. In the United States, a number of states limit the applicability to computer-generated decisions and typically ensure an appeals process where a human makes any final decisions.³³

This ability to interact with other moral agents is thus central to treating others morally. Being in an interpersonal relationship allows all sides to give and take reasons regarding how they are to be treated by the other and to take up relevant factors that they may not have considered beforehand.³⁴ In fact, what might distinguish machine-made legal judgments from human ones is the human ability to establish what is relevant as part of the judicial process rather than in advance. That ability is, in fact, what creates space for sentiments such as mercy and compassion to arise. This point is why only persons—so far, at least—can show respect for other persons.

So, if it seems wrong to subject persons to legal penalties based on machine judgment, it seems even more wrong to subject them to life-and-death decisions based on machine judgment. A machine might be able to enforce the law, but it is less clear if it can



A Stryker vehicle commander interacts in real time with a Soldier avatar that is participating remotely from a collective trainer. The U.S. Army Research Laboratory, University of Southern California Institute for Creative Technologies, Combined Arms Center-Training and Program Executive Office for Simulation, Training and Instrumentation are developing the Synthetic Training Environment, which will link augmented reality with live training. (U.S. Army photo by Lt. Col. Damon "DJ" Durall)

provide justice. Sparrow further observes that what distinguishes murder from justified killing cannot be expressed by a “set of rules that distinguish murder from other forms of killing, but only by its place within a wider network of moral and emotional responses.”³⁵ Rather, combatants must “acknowledge the morally relevant features” that render another person a legitimate target for killing. In doing so, they must also grant the possibility that the other person may have the right not to be attacked by virtue of their noncombatant status or other morally relevant feature.³⁶

The concern here is not whether using robots obscures moral responsibility; rather, it is that the employment of AI systems obscures the good that humans *can* do, even in war. Because humans can experience mercy and compassion, they can choose not to kill, even when, all things being equal, it may be permissible.

The fact that AI-driven systems cannot have the kind of interpersonal relationships necessary for moral behavior accounts, in part, for much of the opposition to their use.³⁷ If it is wrong to treat persons as mere means, then it seems wrong to have a “mere means” be in a position to decide how to treat persons. One problem with this line of argument, which Sparrow recognizes, is that not all employment of autonomous systems breaks the relevant interpersonal relationship. To the extent humans still make the decision to kill or act on the output of a DSS, they maintain respect for the persons affected by those decisions.

However, even with semi-autonomous weapons, some decisionmaking is taken on by the machine—mediating, if not breaking, the interpersonal relationship. Here Scharre’s point is relevant. Morality may demand an interpersonal relationship between killer and killed but as a matter of practice, few persons in those roles directly encounter the other. An Islamic State fighter would have no idea whether the bomb that struck him was the result of

a human or a machine process; therefore, it does not seem to matter much which one it was. A problem remains, however, regarding harm to noncombatants. While, as a practical matter, they have no more experience of an interpersonal relationship than a combatant in most cases, it still seems wrong to subject decisions about *their* lives and deaths to a lethal AI system, just as it would seem wrong to subject decisions about one’s liberty to a legal AI system. Moreover, as the legal analogy suggests, it seems wrong even if the machine judgment were the correct one.

This legal analogy, of course, has its limits. States do not have the same obligations to enemy civilians that they do toward their own. States may be obligated to ensure justice for their citizens but are not so obligated to citizens of other states. There is a difference, however, between promoting justice and avoiding injustice. States may not be obligated to ensure justice for citizens of another state; however, they must still avoid acting unjustly toward them, even in war. So, if states would not employ autonomous weapons on their own territory, then they should not employ them in enemy territory.³⁸

Here, of course, conditions matter. States may not choose not to employ LAWS in their own territory in conditions of peace; however, given the stakes, they may reasonably choose to do so in war, precisely because they are less lethal. If that were to be the case, then the concern regarding the inherent injustice of AI-driven systems could be partially resolved. Of course, it is not enough that a state treats enemy civilians with the same standards it treats its own. States frequently use their own citizens as mere means to an end, so we would want a standard for that treatment that maintained a respect for persons.

For the action of a state to fully meet the standards necessary to count as respecting persons, it must be taken for the sake of all those who may be affected. This condition does not mean that some

may not be harmed; however, if all alternatives would result in the same harm to that person, then it makes sense to choose the alternative that harms the fewest persons. As Isaak Applbaum argues, “If a general principle sometimes is to a person’s advantage and never is to that person’s disadvantage, then actors who are guided by that principle can be understood to act for the sake of that person.”³⁹ So, to the extent AI-driven systems do make targeting more precise than human-driven ones, as well as reducing the likelihood that persons will be killed out of revenge, rage, frustration, or just plain fatigue, then their employment would not put any persons at more risk than if those systems were not employed. To the extent that is the case, then arguably states are at least permitted, if not obligated, to use them. Because employing these systems under such conditions constitutes acting for the sake of those persons, it also counts as a demonstration of respect toward those persons, even if the interpersonal relationship Sparrow described is mediated, if not broken, by the machine.

Moral Hazards of AI Systems

While the use of AI systems may not be inherently evil, the dual concerns of responsibility and dehumanization suggest their use will give rise to a number of moral hazards that need to be addressed if states are to ethically use these systems to their full capability. Moral hazards arise when one person assumes greater risk because they know some other person will bear the burden of that risk.⁴⁰ Given the reduction in risk—both for political leaders who decide to use force as well as the combatants who employ it—the employment of AI systems will establish an incentive structure to ignore the moral risks described above. To address this concern, we need to have a better account of how moral autonomy relates to machine autonomy and how humans, who have moral autonomy, can relate to machines.

Psychological Effects: Desensitization and Trauma

In general, trends in military technology have been to distance soldiers from the killing they do. Crossbows allowed killing at greater distances than did swords, rifles farther than crossbows, cannons and artillery farther than rifles. What is different about autonomous weapons is that they do not just distance soldiers from killing, they can also distance soldiers from the decision to kill. While this is clearly true in fully autonomous systems, it can be true for semi-autonomous systems as well. As P.W. Singer notes in *Wired for War*;

By removing warriors completely from risk and fear, unmanned systems create the first complete break in the ancient connection that defines warriors and their soldierly values.⁴¹

As Singer goes on to observe, the traditional warrior identity arises from conquering profound existential fear, “not the absence of it.”⁴² The result is a fighting force that is not merely distanced from risk, but disconnected from it altogether. As one Air Force lieutenant reportedly said about conducting unmanned airstrikes in Iraq, “It’s like a video game. The ability to kill. It’s like . . . freaking cool.”⁴³

Of course, desensitization is not the only reaction operators have had to the use of autonomous and semi-autonomous systems. In fact, in 2015 a large number of drone operators quit, some citing overwork and others citing the horrors they felt responsible for as reasons.⁴⁴ As Samuel Issacharoff and Richard Pildes observe, the use of LAWS has, in some cases at least, increased the individuation of responsibility for killing and thus brought about a greater sense of responsibility for the killing they do.⁴⁵

One feature that increases sensitivity is the amount of time unmanned aerial vehicle (UAV) pilots spend observing their targets and then watching the effects and aftereffects of strikes they initiate. One U.S. operator, Brandon Bryant,

reported as a source of emotional stress the fact that after a strike, he would not only sometimes have to review the aftermath, but often watch his targets die. Recounting one strike in Afghanistan, he not only observed the strike but also the bodies and body parts afterward. One particularly disturbing image was watching one of the individuals struck. As he recalls, “It took him a long time to die. I just watched him. I watched him become the same colour [sic] as the ground he was lying on.”⁴⁶

This kind of interaction is typically not a feature of conventional strikes. As one UAV pilot put it, “I doubted whether B-17 and B-20 [sic] pilots and bombardiers of World War II agonized much over dropping bombs over Dresden or Berlin as much as I did taking out one measly perp in a car.”⁴⁷ The point here is not that increased use of semi-autonomous and autonomous weapons will bring about more or less sensitivity to killing or trauma but rather that as the character of war changes, different persons will respond differently. The ethical imperative is that leaders pay attention to those changes and take steps to mitigate their ill effects.

Lowering the Threshold to War

Lowering risk to soldiers also lowers risk for civilian leadership when it comes to decisions regarding when to use such weapons. Of course, this concern is not unique to autonomous systems. Any technology that distances soldiers from the violence they do or decreases harm to civilians will lower the political risks associated with using that technology. The ethical concern here is to ensure that decreased risk does not result in an increase in the number of unjust uses of these weapons.⁴⁸ Otherwise, one offsets the moral advantage gained from greater precision.

As Christian Enemark argues, “Political leaders, having less cause to contemplate the prospect of deaths, injuries, and grieving families, might accordingly feel less anxious about using force to solve political problems.”⁴⁹ Like concerns regarding

desensitization, concerns regarding lowering the threshold to war may be overstated. While arguably the use of UAV strikes has expanded over the last decade, instances of escalation into wider conflict have not. Even in areas such as Pakistan, Yemen, and Somalia, where the United States is not at war, the conflict in question preceded the use of unmanned systems, not the other way around. Thus, the ethical question is whether, if this technology were not available, the United States would (and should) do *something*. If the answer is yes, then to the extent the use of force is just and the use of LAWS makes the use of force more precise and humane, it is at least permissible. If the answer is no, then it is likely that no force would be permissible.

The difficulty in resolving this concern is that, much like the concern regarding desensitization, it pits a psychological claim regarding human motivations to employ violence against moral claims associated with the permissibility of violence. The answer to one question is not an answer to the other. Thus, while it may be true that lower risks make decisions about using force easier, it is irrelevant to whether such force is permissible. Having said that, to the extent the psychological concern is valid, it makes sense to confirm that decisions to use risk-decreasing weapons are subject to strict oversight to ensure the conditions of justice are met as well as any other measures that might mitigate these effects. The absence of this oversight and transparency is, in fact, often cited in the literature as a genuine moral concern and has been a longstanding criticism of the U.S. UAV operations.⁵⁰ Given this concern, it makes sense to ensure such oversight and transparency are in place. In this way one can ensure the human reliance on the machine does not set humans up for moral failures they may otherwise not make.

Another concern is that even when LAWS are employed ethically in the service of legitimate U.S. interests, their use may drag the United States into local conflicts of questionable justice. Enemark

notes a debate within DOD regarding whether such strikes are permitted only against high-value targets or also against the larger number of low-level militants whose concerns are more local. He observes that the narrower set is more defensible as preemptive strikes to the extent these individuals are actively plotting against the United States whereas lower-level militants are motivated to fight for local concerns. As he states;

The narrow view is more easily defensible because individuals who are actively plotting to attack the United States more obviously attract (pre-emptive) defensive action than do individuals who merely happen to possess an antipathy towards the United States.⁵¹

Of course, this concern arises as much out of the fact that networks of terrorists threatening the United States draw on and cooperate with networks of oppositionists whose concerns are local, sometimes to the point where it is difficult to distinguish between the two. Thus, regardless of the means used, engaging the former risks expanding conflict with the United States to the latter, who would not otherwise be a threat. While this concern is real, it is more a feature of the character of the conflicts the United States finds itself in rather than the weapons system itself. In fact, it is conceivable that AI-assisted analysis could increase the U.S. military's capability to differentiate between these local and transnational networks. Having said that, the fact of this dynamic suggests the United States should adopt the narrower policy and employ a principle of conservatism when pressure to expand targeting to local targets increases. It may be permissible to do so; however, there should be a demonstrable relationship between the putative target and any threat to the United States.

Automation Bias

One other concern, touched upon earlier, is that humans can sometimes depend too much on the

machine for decisions. One of the most often cited examples of this phenomenon is the shootdown of Iranian Air flight 655 by USS *Vincennes* in 1988. The USS *Vincennes* was equipped with the Aegis ballistic missile defense system, which is fully autonomous but has humans monitoring it as it goes through its targeting cycle. Humans can override the system at any point in this cycle and, in fact, the system was set to its lowest degree of autonomy. The jet's path and radio signature were consistent with civilian airliners; however, the system registered the aircraft as an Iranian F-14 and thus as an enemy. Though the data was telling the crew the aircraft was civilian, they trusted the computer's judgment and shot it down anyway, resulting in the deaths of 290 passengers.⁵³

The difficulty for humans in situations like this is that the complexity of machine "thinking" coupled with the pressure to act, especially in combat, disposes them to trust the machine, especially when doing so can absolve them of at least some of the responsibility of the action in question as well as avoid the consequences of inaction. Moreover, that trust can emerge independent of the reliability of the machine. One study conducted by Korean researchers indicated that the most important factors in human assimilation of DSS were institutional pressure, mature information technology infrastructure, and top management support. Quality of information, stated the report, had no significant impact on DSS assimilation.⁵⁴

Thus, the concern with such systems is that even though humans can prevent wrongful machine behavior, often they will not. That counterintuitive outcome arises from the fact that what the machine often presents to the human is a judgment, but the human takes it as fact. This certainly seemed to be case in the shootdown of the Iranian airliner. The *fact* was that there was an aircraft approaching *Vincennes*, which the system judged was enemy. From the context, specifically the flight path and

radio signature, the humans on board should have questioned the machine and aborted the attack.⁵⁵ As machine judgments become more complex, this concern is only going to be heightened.

This point suggests that operators are going to need to develop sufficient expertise to know what sources of bad judgment are. It will also require operators to adopt a “principle of conservatism” regarding when they should trust the machine without corroborating its output, and limit those times to only what is necessary to accomplishing the mission at hand. To facilitate that trust, as Scharre and Horowitz argue, designers will have to do their best to ensure the outputs of AI systems are “explainable” to at least the operator, if not the commander.⁵⁶

The good news here, as Scharre points out, is that the most successful AI systems will be those that rely on human-machine interaction, suggesting the most successful systems will have a human integrated into the decision cycle.⁵⁷ These systems, which he refers to as “Centaur systems,” are intentionally designed to maximize the speed and accuracy of a hybrid human-machine system in given situations. Examples include defensive systems such as the counter-rocket, artillery, and mortar systems that autonomously create “do not engage” sectors around friendly aircraft. These systems have a human fail-safe to ensure engagements outside those sectors avoid fratricide or harm to civilian aircraft that might approach too closely.⁵⁸

Conclusion

What this analysis has shown is that the arguments for considering military AI systems, even fully autonomous ones, *mala en se* are on shakier ground than those that permit their use. It is possible to reduce, if not close, the responsibility gap and demonstrate respect for persons even in cases where the machine is making all the decisions. This point suggests that it is possible to align effective AI systems development with our moral commitments

and conform to the war convention.

Thus, calls to eliminate or strictly reduce the employment of such weapons are off base. If done right, the development and employment of such weapons can better deter war or, failing that, reduce the harms caused by war. If done wrong, however, these same weapons can encourage militaristic responses when other nonviolent alternatives were available, resulting in atrocities for which no one is accountable, and desensitizing soldiers to the killing they do. To promote the former and avoid the latter, the United States should consider the following measures to ensure the ethical employment of these weapons systems:

- *Work with AI-developing states to update international law.* As previously discussed, international law abhors a vacuum and makes the introduction of any system that mitigates or removes human responsibility problematic. On the other hand, many AI-developing states will take advantage of this vacuum to maximize the effectiveness of these systems, sometimes, if not often, without regard to the moral concerns discussed above. This point suggests the need to update the IHL and other applicable international law, to specify standards of responsibility for the employment of semi-autonomous and fully autonomous systems. These standards would include something like;
 - Operators would have to justify trust in any facts or judgments made by the machine. If that justification is inadequate, they may be responsible for any violations.
 - Operators should ensure AI systems are only employed in conditions for which they are designed to perform ethically. They are also responsible for monitoring the environment and the machine and ceasing operations when conditions changes in a way that sets conditions for violations.

- *Establish standards for diffusing responsibility.* States should establish standards for holding acquisition officials, programmers, designers, and manufacturers responsible for machine violations. These standards will be especially important for fully autonomous systems where conditions for trust by commanders and operators are heavily dependent on the procurement side for ensuring that the machine meets standards associated with operational and functional morality. Meeting such standards would entail accounting for IHL when determining the features of the machine in the same way one might incorporate a safety device on a rifle.
- *Maintain a reasonably high threshold for use.* To ensure employment of AI systems does not inappropriately lower the threshold to violence, states should agree to only employ these systems when the conditions of *jus ad vim* are met. These conditions permit an armed response for acts of aggression that fall short of war, but include the other standards of *jus ad bellum* as well as the requirement to take steps to avoid escalation. *Jus ad vim* also entails an obligation to ensure a high degree of probability that the use of force will achieve the desired objective.⁵⁹
- *Specify conditions for employment.* Given the different human-machine relationships, states should specify conditions for use that ensure meaningful human control and the appropriate trust relationships are maintained. Update these standards as the technology evolves to avoid further gaps between effective use of AI systems and moral commitments.
- *Regulate AI proliferation.* States that develop AI systems for military use should establish proliferation standards similar to the ones the United States has established for the proliferation of UAVs. At a minimum, these standards should include a commitment to only employ these systems in conflicts that meet the standards of

jus ad bellum and in a manner that meets the standards of *jus in bello*. Moreover, there should be a strong presumption of denial to recipients of the technology who have, in the past, been weak on their commitments to these standards.⁶⁰

- *Preserve the soldier identity and address conditions that give rise to desensitization and other psychological trauma.* As the U.S. military becomes more reliant on AI technology, soldiers will experience less risk, but not less trauma. Senior leaders should continue efforts to understand the nature of this trauma and take steps to mitigate it. Moreover, disconnecting soldiers from the risk will also affect how society views and rewards military service. Senior leaders should take steps now to mitigate this potential moral hazard. One step could be to rotate AI system operators in and out of assignments that expose them to risks commensurate with the conflict in question. Doing so will prevent the creation of a class of “riskless” soldiers and moderate the impact of this technology on civil-military relations.
- *Communicate the principles regarding AI use.* Military leaders should develop a communications plan to explain the ethical framework for AI use to the public, media, and Congress.⁶¹

As Sharkey observes, the heavy manpower requirements with remotely controlled systems will place greater pressure to design and employ increasingly autonomous systems.⁶² This point, coupled with the increased effectiveness these systems afford, suggests the trend toward fully autonomous systems is inevitable. As this pressure mounts, commitments to keep humans in the decision process will be increasingly difficult to uphold. Fortunately, as the above analysis indicates, it is possible to manage the moral hazards associated with this technology to ensure moral commitments to human dignity, the rule of law, and a stable international

order are met. Doing so may not assuage every Google employee; however, it will ensure that in acquiring these systems, the United States avoids evil. PRISM

Notes

¹ Scott Shane, Cade Metz, and Daisuke Wakabayashi, “How a Pentagon Contract Became an Identity Crisis for Google,” *New York Times*, May 30, 2018, available at <<https://www.nytimes.com/2018/05/30/technology/google-project-maven-pentagon.html>>. It is worth noting that Google’s commitment to avoiding evil is less than consistent, given its proposed cooperation with the Chinese government to provide a censored search engine. See Rob Schmitz, “Google Plans for a Censored Search Engine in China,” National Public Radio, August 2, 2018, available at <<https://www.npr.org/2018/08/02/635047694/google-plans-for-a-censored-search-engine-in-china>>.

² “Thursday Briefing: EU Calls for Ban on Autonomous Weapons,” *Wired*, September 18, 2018, available at <<https://www.wired.co.uk/article/wired-awake-130918>>.

³ Future of Life Institute, “Lethal Autonomous Weapons Pledge,” available at <<https://futureoflife.org/lethal-autonomous-weapons-pledge/?cn-reloaded=1>>.

⁴ For the purposes of this discussion, the term AI systems will refer to military *AI systems* that may be involved in “life-and-death” decisions. These systems include both lethal autonomous weapons systems that can select and engage targets without human intervention and decision support systems that facilitate complex decisionmaking processes, such as operational and logistics planning.

⁵ Paul Mozur, “Beijing Wants AI to be Made in China by 2030,” *New York Times*, July 20, 2017, available at <<https://www.nytimes.com/2017/07/20/business/china-artificial-intelligence.html>>.

⁶ Erwin Danneels, “Disruptive Technology Reconsidered: A Critique and Research Agenda,” *Journal of Product Innovation Management* 21, no. 4 (July 2004), 21, 249.

⁷ Ronald G. Arkin, *Governing Lethal Behavior in Autonomous Robots* (Boca Raton, FL: Chapman and Hall, 2009), 30.

⁸ Colonel James Boggess, interview with author, January 4, 2017.

⁹ Noel Sharkey, “Killing Made Easy,” in *Robot Ethics*, ed. Patrick Lin, Keith Abney, and George A. Berkey (Cambridge, MA: MIT Press, 2014), 113.

¹⁰ Much of the literature describes human interaction in terms of “humans in the loop,” where the human makes decisions for the machine; “humans on the loop,”

where humans monitor the machine and only act to prevent it from acting error; and “humans off the loop,” where humans provide no additional direction to the machine once it is set in operation. These terms, however, are somewhat imprecise, as in all the three instances, humans played a role in how the machine made a decision, whether in design, production, or employment. Because of this imprecision, I will avoid use of these terms for this discussion.

¹¹ Wendell Wallach and Colin Allen, *Moral Machines: Teaching Robots Right from Wrong* (New York: Oxford University Press, 2009), 16.

¹² Ibid.

¹³ Ibid.

¹⁴ Marcello Guarini and Paul Bello, “Robotic Warfare: Some Challenges in Moving from Noncivilian to Civilian Theaters,” in Lin, Abney, and Berkey, *Robot Ethics*, 129–130.

¹⁵ Ibid., 131–132.

¹⁶ Zed Adams and Jacob Browning, “Introduction,” in *Giving a Damn: Essays in Dialogue with John Haugeland*, ed. Zed Adams and Jacob Browning (Cambridge, MA: MIT Press, 2017), 4.

¹⁷ Kevin Warwick, “Robots with Biological Brains,” in Lin, Abney, and Bekey, *Robot Ethics*, 327–328.

¹⁸ Adams and Browning, 5.

¹⁹ Ibid., 13.

²⁰ Wendell Wallach and Colin Allen, “Framing Robot Arms Control,” *Ethics of Information Technology* 15 (2013), 127.

²¹ Paul Scharre and Michael Horowitz, *Artificial Intelligence: What Every Policy Maker Needs to Know* (Washington, DC: Center for a New American Security, June 2018), 11, available at <<https://www.cnas.org/publications/reports/artificial-intelligence-what-every-policymaker-needs-to-know>>. As an example of “black box” thinking, Scharre and Horowitz note that an “AI image recognition system may be able to identify the image of a school bus, but not be able to explain what features of the image cause it to conclude that the picture is a bus.”

²² Heather Roff, “Killing in War: Responsibility, Liability, and Lethal Autonomous Robots,” in *The Routledge Handbook of Ethics in War*, ed. Fritz Allhoff, Nicholas G. Evans, and Adam Henschke (New York: Routledge, 2013), 355.

²³ Hin Yan Liu, “Refining Responsibility: Differentiating Two Types of Responsibility Issues Raised by Autonomous Weapons Systems,” in *Autonomous Weapon Systems: Law, Ethics, Policy*, ed. Nehal Bhuta et al. (Cambridge: Cambridge University Press, 2016), 341–344.

²⁴ Geoffrey Brennan et al., *Explaining Norms*

(Oxford: Oxford University Press, 2013), 35–39.

²⁵ These principles are also stated in Army doctrine. See Army Doctrine Reference Publication (ADRP) 1, *The Army Profession* (Washington, DC: Headquarters, Department of the Army, June 2015), 3–5. I owe this point to Michael Toler, Center for the Army Profession and Ethic.

²⁶ Noah Sharkey, “Guidelines for the Human Control of Weapon Systems,” International Committee for Robot Arms Control, April 2018, available at <https://www.icrac.net/wp-content/uploads/2018/04/Sharkey_Guideline-for-the-human-control-of-weapons-systems_ICRAC-WP3_GGE-April-2018.pdf>.

²⁷ Scharre and Horowitz, *Autonomy in Weapons Systems*, 16.

²⁸ Geoffrey S. Corn, “Risk, Transparency, and Legal Compliance,” in Bhuta et al., *Autonomous Weapon Systems*, 217–218.

²⁹ Michael Walzer, *Arguing About War* (New Haven: Yale University Press, 2004), 23–32.

³⁰ Paul Scharre, *Army of None: Autonomous Weapons and the Future of War* (New York: W.W. Norton and Company, 2018), 287–288.

³¹ Robert Sparrow, “Robots and Respect: Assessing the Case Against Autonomous Weapon Systems,” *Ethics and International Affairs* 30, no. 1 (2016), 106. It is worth noting here that the Army Ethic acknowledges the importance of respect and includes as a principle, “In war and peace, we recognize the intrinsic dignity and worth of all people, treating them with respect.” See ADRP 1, 2–7. I owe this point to Michael Toler, Center for the Army Profession and Ethic.

³² Eliav Liebllich and Eyal Benvenisti, “The Obligation to Exercise Discretion in Warfare: Why Autonomous Weapons Systems Are Unlawful,” in Bhuta et al., *Autonomous Weapon Systems*, 266.

³³ *Ibid.*, 267.

³⁴ The point here is not that the battlefield is a place to negotiate or that there has to be some kind of interaction independent of a targeting process to justify the decision to use lethal force. Rather, the point is that in any given situation where a human being may be harmed that the decision made to commit that harm is made by another human who can identify and consider the range of factors that would justify that harm. In this way, the person deciding to harm understands it is another person whom he or she is harming and considers reasons not to do it. Autonomous systems may eventually be able to discern a number of relevant factors; however, that only entails they are considering reasons to harm, not reasons not to harm.

³⁵ Sparrow, “Robots and Respect,” 101.

³⁶ *Ibid.*, 106–107.

³⁷ *Ibid.*, 108.

³⁸ Liebllich and Benvenisti, “The Obligation to Exercise Discretion in Warfare,” 267.

³⁹ Arthur Isaak Applbaum, *Ethics for Adversaries: The Morality of Roles in Public and Professional Life* (Princeton, NJ: Princeton University Press, 1999), 162–166. Applbaum refers to situations where someone is better off and no one is worse off as “avoiding Pareto-inferior outcomes.” Avoiding such outcomes can count as “fair” and warrant overriding consent.

⁴⁰ Kenneth J. Arrow, “Uncertainty and the Welfare Economics of Medical Care,” *American Economic Review* 53, no. 5 (December 1963), 941, 961. See also Matthew McCaffrey, “Moral Hazard: Kenneth Arrow vs Frank Knight and the Austrians,” MISES WIRE, March 14, 2017, available at <<https://mises.org/blog/moral-hazard-kenneth-arrow-vs-frank-knight-and-austrians>>.

⁴¹ P.W. Singer, *Wired for War: The Robotics Revolution and Conflict in the 21st Century* (New York: Penguin Books, 2009), 332.

⁴² *Ibid.*

⁴³ *Ibid.*, 395.

⁴⁴ Pratap Chatterjee, “American Drone Operators Are Quitting in Record Numbers,” *The Nation*, March 5, 2015, available at <<https://www.thenation.com/article/american-drone-operators-are-quitting-record-numbers/>>.

⁴⁵ Samuel Issacharoff and Richard Pildes, “Drones and the Dilemmas of Modern Warfare,” in *Drone Wars: Transforming Conflict, Law, and Policy*, ed. Peter Bergen and Daniel Rothenberg (Cambridge: Cambridge University Press, 2015), 399.

⁴⁶ Phillip Sherwell, “U.S. Drone Pilot Haunted by Horrors of Remote Killings; Operator Sickened by Death Count,” *London Daily Telegraph*, October 25, 2013, available at <<https://www.telegraph.co.uk/news/worldnews/northamerica/usa/10403313/Confessions-of-a-US-drone-operator-I-watched-him-die-It-took-a-long-time.html>>.

⁴⁷ Matt J. Martin and Charles Sasser, *Predator: The Remote-Control Air War over Iraq and Afghanistan: A Pilot’s Story* (Minneapolis, MN: Zenith Press, 2010), quoted in Issacharoff and Pildes, 416.

⁴⁸ Christian Enemark, *Armed Drones and the Ethics of War* (London: Routledge, 2014), 22.

⁴⁹ *Ibid.*, 23.

⁵⁰ Sharkey, “Guidelines for the Human Control of Weapon Systems,” 115. As Sharkey states, “It is now unclear what type and level of evidence is being used to sentence nonstate actors to death by Hellfire attack without right to appeal or right to surrender.” The point here is not that U.S. targeting procedures are unjust, but to the extent they are not transparent—at least to the

extent possible without compromising the process—the perception of injustice will persist, generating the kind of response evidenced by the Google employees.

⁵¹ Enemark, *Armed Drones and the Ethics of War*, 25.

⁵² I owe this point to Colonel David Barnes, Department of English and Philosophy, U.S. Military Academy at West Point.

⁵³ Singer, *Wired for War*, 125.

⁵⁴ Hyun-Ku Lee and Hangjung Zo, “Assimilation of Military Group Decision Support Systems in Korea: The Mediating Role of Structural Appropriation,” *Information Development* 21, no. 1 (2017), quoted in James Boggess, “More than a Game: Third Offset and the Implications for Moral Injury,” in *Closer than You Think: The Implications of the Third Offset Strategy* (Carlisle, PA: Strategic Studies Institute, 2017), 133.

⁵⁵ David Evans, “Vincennes: A Case Study,” *Proceedings* 119, no. 8/1086 (August 1993).

⁵⁶ Scharre and Horowitz, *What Policy Makers Need to Know*, 11.

⁵⁷ Scharre, *Army of None*, 321.

⁵⁸ *Ibid.*, 323–324.

⁵⁹ Daniel Brunstetter and Megan Braun, “From Jus ad Bellum to Jus ad Vim: Recalibrating our Understanding of the Moral Use of Force,” *Ethics and International Affairs* 27, no. 1 (2013).

⁶⁰ Fact Sheet, “U.S. Policy on the Export of Unmanned Aerial Systems,” U.S. State Department, April 19, 2018, available at <<https://www.state.gov/r/pa/prs/ps/2018/04/280619.htm>>. This policy is a little more permissive than the one implemented by the last administration. It permits the sale of armed unmanned systems through direct commercial sales and removes language regarding a “strong presumption of denial” for systems that cross the Missile Technology Control Regime thresholds, which in this case are systems with a range greater than 300 kilometers and are capable of carrying payloads of 500 kilograms or more. Both sets of standards require upholding international law. See Fact Sheet, “U.S. Policy on the Export of Unmanned Aerial Systems,” U.S. State Department, February 17, 2015, available at <<https://2009-2017.state.gov/r/pa/prs/ps/2015/02/237541.htm>>.

⁶¹ I owe this point to Dr. Steve Metz, Strategic Studies Institute, U.S. Army War College.

⁶² Sharkey, “Guidelines for the Human Control of Weapon Systems,” 115.