

6-2013

The Future of Military Virtue: Autonomous Systems and the Moral Deskilling of the Military

Shannon Vallor

Santa Clara University, svallor@scu.edu

Follow this and additional works at: <http://scholarcommons.scu.edu/phi>



Part of the [Ethics and Political Philosophy Commons](#)

Recommended Citation

Vallor, Shannon. "The Future of Military Virtue: Autonomous Systems and Moral Deskilling in the Military Profession." In 2013 5th International Conference on Cyber Conflict (CyCon 2013): Proceedings, Karlis Podens, Jan Stinissen and Markus Maybaum, eds. (Tallinn, Estonia: NATO CCDCOE, 2013), pp. 471-486."

Copyright © 2013. Reprinted with permission. https://ccdcoe.org/cycon/2013/proceedings/d2r1s10_vallor.pdf

This Book Chapter is brought to you for free and open access by the College of Arts & Sciences at Scholar Commons. It has been accepted for inclusion in Philosophy by an authorized administrator of Scholar Commons. For more information, please contact rscroggin@scu.edu.

Permission to make digital or hard copies of this publication for internal use within NATO and for personal or educational use when for non-profit or non-commercial purposes is granted providing that copies bear this notice and a full citation on the first page. Any other reproduction or transmission requires prior written permission by NATO CCD COE.

“The Future of Military Virtue: Autonomous Systems and Moral Deskillling in the Military Profession.” In 2013 5th International Conference on Cyber Conflict (CyCon 2013): Proceedings, Karlis Podens, Jan Stinissen and Markus Maybaum, eds. (Tallinn, Estonia: NATO CCDCOE, 2013), pp. 471-486.

https://ccdcoe.org/cycon/2013/proceedings/d2r1s10_vallor.pdf

The Future of Military Virtue: Autonomous Systems and the Moral Deskilling of the Military

Shannon Vallor

Department of Philosophy
Santa Clara University
Santa Clara, California USA
svallor@scu.edu

Abstract: Autonomous systems, including unmanned aerial vehicles (UAVs), anti-munitions systems, armed robots, cyber attack and cyber defense systems, are projected to become the centerpiece of 21st century military and counter-terrorism operations. This trend has challenged legal experts, policymakers and military ethicists to make sense of these developments within existing normative frameworks of international law and just war theory. This paper highlights a different yet equally profound ethical challenge: understanding how this trend may lead to a *moral deskilling* of the military profession, potentially destabilizing traditional norms of military virtue and their power to motivate ethical restraint in the conduct of war. Employing the normative framework of virtue ethics, I argue that professional ideals of military virtue such as courage, integrity, honor and compassion help to distinguish legitimate uses of military force from amoral, criminal or mercenary violence, while also preserving the conception of moral community needed to secure a meaningful peace in war's aftermath. The cultivation of these virtues in a human being, however, presupposes repeated practice and development of skills of moral analysis, deliberation and action, especially in the ethical use of force. As in the historical deskilling of other professions, human practices critical to cultivating these skills can be made redundant by autonomous or semi-autonomous machines, with a resulting devaluation and/or loss of these skills and the virtues they facilitate. This paper explores the circumstances under which automated methods of warfare, including automated weapons and cyber systems, could lead to a dangerous 'moral deskilling' of the military profession. I point out that this deskilling remains a significant risk even with a commitment to 'human on the loop' protocols. I conclude by summarizing the potentially deleterious consequences of such an outcome, and reflecting on possible strategies for its prevention.

Keywords: *automated methods, ethics, military virtue, professionalism, moral deskilling.*

1. INTRODUCTION

Few images highlight the increasingly automated nature of modern warfare better than a photograph of the eerily opaque, windowless nose of the MQ-1 Predator drone, manufactured by General Atomics Aeronautical Systems and a centerpiece of U.S. military and counterterrorism efforts in the Middle East and Africa, where hundreds of targeted drone strikes are now launched annually. Yet drone warfare is merely the leading edge of a broader worldwide trend toward more autonomous methods of warfighting. From South Korea's armed sentry robots guarding the DMZ, to Israel's 'Iron Dome' anti-munitions defense, to miniaturized lethal drones like the U.S. Army's Switchblade, to long-range intercontinental drones like the U.K. Taranis and the U.S. X47-B, militaries around the world are investing in an increasingly automated future. Nor are such investments limited to weapons in the conventional sense. Military and intelligence agencies worldwide are developing increasingly sophisticated and autonomous software algorithms for use in cyberwarfare – conflicts between electronic agents in electronic space that nevertheless have the potential to inflict considerable human losses. Merging with both trends are advancements in algorithms for analysing massive datasets, which can potentially outperform human calculations of threat potential, target value, operational risk, mission cost, casualty estimates and other key strategic variables. Taken together, these developments represent a profound shift in our traditional understanding of the role of human beings in the conduct of war. In this paper I explore one of this shift's most challenging ethical implications, namely, the risk of a significant *moral deskilling* of professional militaries, and a destabilization of associated ideals of military *virtue*.

The broader legal and ethical implications of this shift are immense in scope; they range from the compliance or non-compliance of automated warfighting systems with the laws of war and requirements of just war theory (Asaro 2008), to problems of moral and legal accountability for actions taken by autonomous agents (Sparrow 2007), to the concern that automated methods of warfare are leading to greater 'moral disengagement' among soldiers (Sharkey 2010; Royakkers and van Est 2010). Together, these concerns mandate extensive and widespread critical inquiry and reflection on the automation of war; fortunately, this conversation is now well underway. In addition to scholarly articles by ethicists and legal experts, recent years have seen several high-profile books on related topics (Singer 2009; Arkin 2009; Krishnan 2009; Lin et. al. 2012). Major media outlets from *The New York Times* and *The Wall Street Journal* to online magazines like *Wired* and *Slate* regularly cover emerging developments in automated war technology and their political, legal and ethical ramifications. We are seeing the welcome emergence of a vigorous scholarly and public discourse on the legality and ethics of automated warfare, one likely to

continue to evolve for decades as the possibilities, risks and benefits of automated systems become clearer.

Yet one important subset of these concerns is likely to be less visible to public and political interests than the legalistic and utilitarian dilemmas presently driving the global conversation about the ethics of automated warfare. The subject to which I wish to call attention concerns the future of military virtue in an era of increasingly automated military action. My claim is that unless we take steps to secure that future, we face the possibility of a dangerous *moral deskilling* of the military profession. In what follows, I call the reader's attention to the importance of habitual moral practice and skill acquisition as a precondition for the cultivation of military virtues, which are in turn critical to the standing of militaries as professional bodies with a distinct moral status recognized by combatants and civilians alike. I argue that by depriving soldiers of the opportunity to practice and develop the skills critical to the ethical use of military aggression, increasingly automated methods of warfighting could endanger the moral and professional status of the military vocation. In my conclusion I offer some speculations about how this outcome might be prevented.

2. MORAL PRACTICE AND MILITARY VIRTUE

Before I develop and support my claims, let me briefly explain what 'virtue' in the phrase 'military virtue' entails. The concept of virtue is rooted in classical traditions going as far back as the ancient Greek philosophies of Plato and Aristotle and, in the East, Confucian and Buddhist ethics. It endures today in the writings of contemporary virtue ethicists like Rosalind Hursthouse, Alasdair MacIntyre and Martha Nussbaum, and has found its way into various applied and professional codes of ethics, including business ethics, medical ethics, environmental and engineering ethics (Axtell and Olson 2012). Virtues are habituated states of a person's character that reliably dispose their holders to excel in specific contexts of action, and to live well generally; so moral virtues are states of character that, once acquired, dispose their possessors to perform excellent moral actions of particular sorts, and, more broadly, to excel in moral living. Cardinal examples of moral virtues include wisdom, honesty, courage and moderation; others commonly recognized include loyalty, integrity, respect, honor, patience, compassion and benevolence, though this list is far from exhaustive. How particular virtues are defined and prioritized varies among cultures, historical periods and social roles; yet there is substantial overlap or convergence among diverse virtue traditions, indicating that the qualities seen as most supportive of human flourishing are, while not entirely universal, rooted in widely shared or similar human practices.

Because virtues are habituated rather than inborn, whether or not a person develops

a particular virtue will largely depend on whether they engage repeatedly in the kinds of practices that cultivate it. The virtue of honesty, for example, can only be acquired through repeated practice of truth-telling. Initially, such practice requires guidance by a virtuous model, e.g., someone who is already honest. Over time, repeated practice can lead a person to see for themselves what honesty is, to see it as good in itself and to embody it better and more easily; a person who has cultivated the virtue of honesty is not only consistently inclined to tell the truth, they have learned how to excel at truth-telling in any situation that might arise: who to tell the truth to, when and where, in what way, and to what extent. Moral virtue thus requires more than good will and a steady desire to do the right thing – it requires the cultivation of a kind of *practical wisdom* that directs this right desire intelligently, perceiving and quickly adapting to the unique moral demands of each situation. In his *Nicomachean Ethics*, Aristotle named this practical wisdom *phronesis*; a sort of ‘über-virtue’ that orchestrates one’s individual qualities of moral excellence and integrates them within a complete and flourishing life (1984, 1140a25-30;1145a). The concept of virtue, then, picks out those aspects of persons that enable them to live as moral exemplars for others, qualities of character that we ourselves can strive to cultivate through the same sorts of repeated practice.

The role of virtue in military ethics has long been recognized, and a rich body of existing literature details the way in which virtues like courage, duty, integrity, honor, loyalty and service have historically been inseparable from the ideal of the good soldier (Olsthoorn 2011; Robinson 2007; Reichberg, Syse and Begby 2006; French 2003; Toner 2000). This does not mean that the enterprise of war itself can or should be seen as virtuous. Rather, ideals of virtuous military character, when exercised as normative *expectations* (not just indicators of supererogatory or heroic performance), express a society’s unwillingness to wholly exclude its warfighters from the broader responsibilities and benefits of the moral community.¹ As I argue elsewhere (Vallor 2013), ideals of military virtue, when embedded in the practice and professional identity of military bodies, block the cultural displacement of war to an extra-moral realm where its conduct would be indistinguishable from criminal or mercenary violence.

The ideal serves as a kind of contract between warfighters and the larger community, and when in force, it offers considerable benefits to soldiers and civilians alike. In addition to motivating restraint on the part of soldiers in inflicting civilian harms, it can motivate limited restraint between enemy combatants when each recognizes the other as a professional fighting with honor and moral purpose. It can also support

¹ The term ‘moral community’ here is left deliberately ambiguous in its reference; ideals of military virtue can be seen as ties that bind a soldier to the ethical life of her own nation or culture, or, in more cosmopolitan views, to the ethical life of the global human community of which she is a part.

the psychological integrity of soldiers themselves, by providing a moral context for what are, taken in themselves, brutal and deplorable actions. Finally, it preserves the sense of moral community between warfighters and civilians that allows returning soldiers to be welcomed home, and even valorized. To see the importance of this contingency, one need only be familiar with the starkly different experiences in the United States of veterans of World War II, treated to grand welcoming parades and to this day labelled “The Greatest Generation,” and veterans of the Vietnam War, who returned home to a largely indifferent and often hostile society no longer able to contextualize their service as virtuous.

Thus while war itself cannot be virtuous, because it characteristically impedes rather than supports human flourishing, humans who take on the burdens of military service *can* be - insofar as they aspire to fight only in the manner of an excellent human being. Of course, moral virtue is expressed differently according to the demands of particular circumstances; what compassion and courage call for in battle looks very different from what these demand in civic life. In writing on war and its apparent incompatibility with moral norms, Augustine wrote that precisely because war foments evil (“the desire for harming, the cruelty of revenge, the restless and implacable mind, the savageness of revolting, the lust for dominating”), it is all the more essential that soldiers cultivate the virtuous dispositions of compassion and benevolence to accompany them in battle, so that the “mutual bond of piety and justice” that constitutes common morality has not been irrevocably destroyed by the time that material conditions for peace return (Augustine 1994, 221-222).

Military virtue, then, imperfect as its professional cultivation and practice may be, functions to keep warfighters morally continuous with society. It allows us to see ourselves, and the other, as *worthy* of membership in a moral community, even when engaged in conduct that is in itself destructive to moral community. When the professional cultivation of military virtue is not attempted or its aspirations are abandoned, as in Cambodia, Rwanda, and Srebrenica, the aftermath of war is often precisely what we would expect from Augustine’s account: a shallow peace poisoned by deep distrust, resentment and fear lasting for generations. Survivors of such a moral calamity do not stop suffering when the bloodshed stops: they are burdened with the crippling social degradation that comes from the death of civic norms of trust, mercy, forgiveness, justice and goodwill. Such norms, once destroyed, are not easily reborn; while military virtue may not be able to shield them from assault, it can keep them on life support. For all of these reasons, then, it is essential to the mitigation of the harms of war that military virtue be preserved, both as a meaningful moral concept and as a practical and attainable commitment to ethical warfighting. In what follows I explain why the increasing automation of warfighting methods may jeopardize this imperative.

3. AUTOMATED WARFARE, VIRTUE AND THE MORAL DESKILLING OF MILITARY PRACTICE

Having offered reasons to take the concept of military virtue seriously, I turn to the primary burden of my argument: to show how increasingly automated methods of warfighting challenge the future of military virtue by potentially contributing to a *moral deskilling* of the military profession.

First, let us consider the link between virtues and skills. Aristotle was clear that virtues and skills share many common features – both are acquired by habit and practice, both must be guided by intelligence, and both must be adapted to the demands of given situations. But he also reminds us that virtue is *more* than just skill or know-how; it is a state in which that know-how is reliably put into action when called for, and is done with the appropriate moral concern for what is good: “The agent also must be in a certain condition when he does [virtuous acts]; in the first place he must have knowledge, secondly he must choose the acts, and choose them for their own sakes and thirdly his actions must proceed from a firm and unchangeable character” (1984, 1105a30-35). Someone could have moral skills in the sense of practical moral knowledge but fail to be virtuous because they are unreliable in acting upon this knowledge, or because they act well only for non-moral reasons. Still, moral skills are a *necessary* if not a sufficient condition for moral virtue. Without the requisite cultivation of moral knowledge and skill, even a person who sincerely wishes to do well consistently and for its own sake will be unsuccessful. It follows that if the advancing automation of military conflict were to bring about a significant ‘moral deskilling’ of the profession, the future of military virtue would be gravely endangered.

What would a ‘moral deskilling’ of the military profession amount to, and how might the advancing automation of warfighting systems contribute to it? *Deskilling* is a familiar concept in the analysis of the social impact of technology; for example, we might think of the way in which the skills of machinists and other classes of mechanical labor were devalued by widespread factory adoption of automated machine tools (Braverman 1974). Or consider the worry that the professional work of highly skilled nurses is increasingly given over to a combination of less skilled aides and advanced medical monitoring and medication delivery technologies (Rinard 1996). However, the concept of deskilling has declined in academic usage in the last few decades, in large part because unlike the earlier automation of factory work, the information revolution has thus far seemed to deliver as much *upskilling* as deskilling – workers in many industries have been freed by computers to shift their duties from mindless tasks like filing, copying and collating to more

challenging and knowledge-laden responsibilities. Yet some new applications in information technology may warrant renewed concerns about deskilling, including moral deskilling (Manders-Huits 2006). Whether any automated technology produces deskilling, then, is an empirical question that depends upon the particular context of use. Let us look more closely at the critical meaning of the concept, and how it may apply to the context of automated warfare.

The concept of deskilling has at least two critical implications. The first and most commonly discussed implication is that the deskilling of a given profession may decrease the socioeconomic value, autonomy and power enjoyed by workers, potentially causing them significant psychological and economic harm. A second critical implication, the one I wish to highlight, is that at least for some professions, we may have reason to regret the loss of the professional context for cultivating the given skills because we think the skills themselves are intrinsically valuable. For example, many have mourned the declining skills of artistic handicraft lost to mass manufacture of ready-made objects (Roberts 2010), resulting in renewed interest in ‘handmade,’ ‘custom’ or ‘artisanal’ products. In this context, it is not only the economic welfare of the artisan that we value, and not only the quality of the end product, but also the connection between an artifact and a human whose artistic excellence and knowledge was responsible for its production. We think that it is *good* that humans are skilled at making beautiful and useful objects for their own living, and that even if machines could produce all such goods for us, it would be sad and regrettable if humans were no longer capable of doing the same.

I suggest that the intrinsic value of artisanal skills is not only paralleled but dwarfed by the intrinsic value of moral skills. The concept of *moral* deskilling is only rarely employed used in the sociological literature on technology, in part because sociologists tend to shy away from normative judgments of ethics, and in part because concerns about moral deskilling are sometimes associated with reactionary ‘moral panics’ in reaction to technological change – for example, worries in the 1920’s that the telephone would result in crippling social isolation and the unravelling of people’s capacities for moral interaction. However, the concept remains meaningful, and I suggest that it may have profound significance with respect to the professional impact of military automation. While deskilling has been recognized with respect to the threatened obsolescence of abilities such as those cultivated by military snipers (Townsend and Charles 2008), the more worrisome possibility of a *moral* deskilling of the military profession has yet to be widely acknowledged.

Consider the parallel drawn earlier with artisanal skills. Just as the widespread loss of such skills by humans would not be fully expunged by machines that produce comparable products, a widespread loss of moral skills in the context of military

conduct would not be rendered insignificant by the emergence of machines that produce equivalent, or arguably even better outcomes. This fact has unfortunately been lost in the otherwise rich debate about the legal and ethical implications of automated warfare. A world in which humans involved in warfighting are no longer skilled in the *moral* conduct of war is a world in which the concept of ‘military virtue’ has no meaning. As I have argued elsewhere (2013), where this concept has lost its meaning, the recognition of soldiers as professionals devoted to the selfless service of the moral community is no longer possible.

A. *AUTOMATED WEAPONS SYSTEMS AND MORAL DESKILLING*

Methods of automated warfare may be divided between those involving cyber-conflict, and those involving (directly or indirectly) the application of military force. Let’s first consider the latter, starting with the ongoing debate about autonomous and semi-autonomous weapons. Most of the literature on this subject has focused on the inability of weaponized robots and drone aircraft to act with the moral knowledge, restraint, compassion, discrimination, proportionality and accountability demanded by modern laws of war. A recent Human Rights Watch report on the topic states that their primary concern is the price civilians will pay for these inevitable ethical shortcomings of autonomous weapons systems (Human Rights Watch 2012, 1). But not only is this open to challenge from those with more optimistic projections for artificial moral intelligence (Arkin 2009), it remains silent on the human cost of no longer asking soldiers to cultivate and reliably exercise the same moral capacities. If the optimistic predictions of roboticists are anywhere near correct, we may be moving toward a future where humans start wars, oversee them, and suffer from them, but are no longer *fighting* them, in the concrete sense of making informed and morally reflective choices about who or what gets targeted, or when, in which circumstances, or with what degree of force. My claim is that there is a price to pay here *even if* civilians do not suffer more direct harm as a result.

Consider that the skill set for supervising, approving or vetoing the decisions of semi-autonomous robots seeking to apply lethal force will be much narrower than the skill set required for humans to make those moral decisions themselves. One reason involves the time constraints under which human supervisors of autonomous or semi-autonomous weapons will operate. Many scholars believe that the much-touted principle of humans staying ‘on the loop,’ with veto power over system targeting or firing actions, will soon be rendered largely meaningless when the human operator is given only a fraction of a second to make the veto decision, as is the case with several systems already in operation (Human Rights Watch 2012). One of the key tactical advantages of autonomous weapons systems is that they can make

and execute decisions far faster than humans can. These narrowing time horizons will likely preclude human operators from conducting substantive investigation of, or careful reflection upon, the morally salient features on the ground warranting the robotic application of force. This has already been acknowledged as a fundamental technical obstacle to humans remaining on the loop of engagements between unmanned combat fighters; the delay time injected by satellite communications is simply incompatible with the timescale of air combat (Sharkey 2012).

Add to this the likelihood that human operators of semi-autonomous systems will be tasked with supervising dozens or hundreds of drone or robotic agents at one time, as described by the Swarms.org website for the U.S. military's SWARMS initiative (Scalable sWarms of Autonomous Robots and Mobile Sensors), and the potential for moral deskilling becomes even more evident. We might be tempted to envision human supervisors of autonomous or semi-autonomous weapons systems as elite military judges chosen for their Solomonic wisdom and discretion in the ethical use of lethal force; but in reality they may have even less room for discretion and fewer degrees of decision freedom than air traffic controllers. What sort of moral skill set can we reasonably expect such practices to cultivate? And if moral skills in the use of military force are not cultivated at the level where force is applied, or even at the level where its application is being directly supervised, where *will* it be cultivated, and through what practices?

That advances in automation will result in revolutionary shifts in the skills needed for modern warfighting is news to hardly anyone. In envisioning a future where thanks to advancing automation, "systems and equipment can deploy forward with little if any human presence unless required for acceptance," the U.S. Air Force's published "Flight Plan 2009-2047" for unmanned aerial systems acknowledges that "a key challenge to realizing the vision will be to develop and maintain the right skill sets of systems and operational software developers, mission directors and USAF leaders...leaders will also require different skills to employ air power that is largely non-human" (USAF 2009, 51). But nowhere in this plan is it acknowledged that military leaders traditionally are expected to exercise *moral* as well as technical and strategic skill in the use of weaponized systems; it is worth asking *where* those moral skills will be cultivated in a future of automated warfighting where "relatively few mission directors will be needed" and the skills needed to "prepare, launch and perform" combat operations have been shifted from the field of action to "technology development offices" (*Ibid.*).

One might object by pointing out that decisions to use military force are very rarely conducted under conditions conducive to deep moral reflection. Human soldiers already have to make snap judgments in the field under highly demanding constraints, and even those decisions that can be reviewed by commanders are

rarely evaluated under ideal conditions. Yet it remains the case that a remarkable amount of moral knowledge and skill is presupposed by the human ability to keep a military operation involving lethal force from descending into utter moral chaos. Not every commanding officer has this kind of knowledge and skill, or even adequate exposure to the practices needed for its cultivation; but if *no one* in the chain of command has it, the chances of moral catastrophe are greatly increased. Neither sound rules of engagement nor advance commitments to ‘human values’ will prevent disaster if there is no one who can *apply* them in morally expert ways.

Indeed, moral virtue entails precisely the kind of expertise that allows us to quickly perceive the right course of action even in unpredictable or rapidly changing circumstances, without laborious calculations or clumsy recourse to formal principles. Yet according to many roboticists, a chief advantage of future autonomous weapons systems is that once programmed, they will not need human experts to tell them how to avoid morally catastrophic uses of force, begging the question of whether human soldiers will still be expected to cultivate that expertise for themselves. Remember that repeated moral practice is essential to the cultivation of moral virtue. How might our moral development suffer from transferring the most critical of those practices to machines, whose response times and cognitive architecture will be sufficiently unlike our own to prevent them from serving as models of virtue for us?

Furthermore, if the advance of autonomous weaponry were to lead to a significant ‘moral deskilling’ of the military profession, how would that impact the cultivation of military virtue, which as I claimed earlier, performs a critical function in mitigating the tendency of wars to produce lasting civic devastation? For the sake of argument, let us assume with roboticists like Ron Arkin (2009) that the most optimistic predictions regarding the emergence of artificial moral intelligence will be realized, and that in the not so distant future, human soldiers are no longer regularly called upon to judge when lethal military force is warranted and when it is not, or how it should be applied. We have handed over these judgments to robotic systems without any of our defective dispositions to anger, vengeance, bias, fear and laziness, and with computational abilities that keep their margins of error well below the best-trained of human soldiers. Without opportunities to exercise the skilled moral judgments that the expert application of lethal force requires, what level of ‘moral deskilling’ of professional soldiers may result, and with what consequences for the cultivation of military courage, honor or compassion?

One might interrupt to remind me that the United States has issued a new policy directive ensuring that, as the title of *Wired* magazine’s coverage ably summed up, ‘A Human Will Always Decide When a Robot Kills You’ (Ackerman 2012). This new commitment to reject fully autonomous targeting and lethal engagement should not preclude us from seriously entertaining our thought experiment for

three reasons. First, this directive expires in ten years (Carter 2012), well before most scholars expect reliable technology for fully autonomous lethal robots to be available. Second, it binds only U.S. armed forces, and in no way precludes the development of such systems for other markets. Third, it does not change the above-noted fact that even semi-autonomous weaponry is rapidly shrinking the window for decisions on target selection and engagement below the timescale of human decision-making. Add to this the recognized ‘automation bias’ that leads humans to trust computer judgments over their own (Asaro 2009, 22), and the United States’ promise to preserve human control over lethal means of warfighting may seem less meaningful. Well, then, so much the worse for underperforming, unpredictable and irrational human soldiers, and so much the better for programmable, precise and obedient killer robots, say roboticists like Arkin (2010). They may be right, and from a consequentialist point of view, there is no question that reductions in civilian casualties as a result of increased precision and reduced error rates, if realized, will have to factor into any moral assessment of the use of autonomous lethal weapons.

But the consequentialist equation cannot be the whole story. We must also consider the value of the moral skills that make military virtue and professionalism possible and what their loss might mean for professional soldiers and civilians alike. Before offering some concluding thoughts on how their loss might be prevented, let us move beyond automated weapons systems of the traditional sort, and extend our inquiry to automated methods of cyber-conflict.

B. BEYOND AUTONOMOUS WEAPONS: ALGORITHMIC AUTONOMY, CYBER-CONFLICT AND MORAL DESKILLING

The role of automated systems in warfighting is not limited to drones, robots or autonomous defense munitions. The software algorithms that enable autonomous or semi-autonomous operation of such systems can also be used to automate military or intelligence decision processes that may or may not involve the deployment of autonomous weapons. Consider, for example, a recent paper on an algorithm developed at West Point’s Network Science Center and funded by the U.S. Army Research Office for potential use in ranking the most valuable targets in a terrorist network (Shakarian et. al. 2012). The paper’s authors suggest that the performance of the algorithm, which tends to select mid-level lieutenants in a terrorist network as more valuable targets than high-level commanders, may be superior to independent human assessments of target value. Granting targeting authority to such an algorithm could lead to an operation involving an automated drone strike on the target, but it could also motivate a Special Forces assault or attempt at capture. Yet even without the use of automated weapons, such an operation would embody the trend toward automated warfighting. The moral implications of letting a computer program decide which individual humans deserve to be military targets are starkly

apparent, but again, setting aside obvious worries about the *justice* of automating such a decision, consider its additional implications for human cultivation of moral skills, knowledge and virtue. Imagine that we come to rely upon algorithms of this kind for military and intelligence targeting, but also for determining, for example, whether killing or capturing a particular target is more ethical and prudent, and the best operational design, occasion or ordnance for doing so. What kind of moral character would be required for military officers and other personnel to successfully support such an operation? Would any moral skills or qualities of note be required? If so, what would they be? In what actions would they be cultivated, or exercised? Or would soldiers and mission leaders be called upon strictly as technical specialists, tasked and trusted with nothing more than ensuring informational integrity in the communication of algorithmic decisions down the chain of command?

Such questions clearly extend into the realm of military operations involving no direct deployment of force whatsoever, such as cyber-warfare, or more broadly, cyber-conflict. Consider an algorithm that is programmed to defend government networks from intrusion, and to launch a counter-attack upon any electronic system or network it identifies as the host of the intruding informational agent. Set aside for now the technical questions about how to effectively design such an algorithm, such that it does not frequently mistake benign interactions for a cyberattack, or misidentify the agent responsible for an attack. To run parallel with our thought experiment about autonomous robots and other weapons systems, let us assume, just for the sake of argument, that we will soon develop sufficiently advanced artificial intelligence such that we can trust such algorithms to select, as justice requires, a proportional and discriminating response to any given cyberattack. What skills and virtues would be required of the human operators and supervisors of a cyberdefense system driven by such algorithms?

Let us say that we adopt the policy that such systems must maintain a human ‘on the loop,’ who in each case is tasked with approving or rejecting the system’s request to launch a counterattack. Even setting aside the ‘automation bias’ mentioned above, which can already predispose us to defer to computer decisions (Asaro 2009), how would such a supervisor ever become qualified to make that judgment, in a professional setting where the decision process under review is no longer regularly exercised by humans in the first place? An expert supervisor of another’s decision, in order to be worthy of the authority to override it, must have acquired expertise in making decisions of the very same or a similar kind. The requisite skills and wisdom that constitute such expertise could only be acquired, according to most theories of expertise, by having repeatedly and habitually practiced the actions in question, with an opportunity to learn from mistakes and successes, and to receive corrective feedback from others who already have the expertise one seeks to acquire. Where will the human supervisors of automated cyber-conflict acquire such practice, and the expertise in the proportionate and discriminating use of cyber-power that it

alone can engender? And what are the implications for human beings engaged in cyber-conflict, and for those impacted by cyber-conflicts, if they do not?

4. CONCLUSIONS

Of course, the reduction of human decision-making to mechanistic, formulaic or quasi-algorithmic processes can happen by means other than technological automation. We can easily conceive of military environments in which soldiers and officers are encouraged to eschew complex moral reasoning in favor of legalistic templates, decision-trees and other formal mechanisms of reducing the cognitive burdens (and freedoms) of human judgment.² Thus any ‘moral deskilling’ of the military profession need not be *essentially* linked to advances in the technological automation of war – it may have other causes as well. That said, the considerations above make it clear that advances in technological automation may greatly exacerbate any existing defects in the ability of today’s military bodies to cultivate moral skills and virtues among their members and within their leadership ranks. What can be done to prevent such an eventuality? One option, of course, is a wholesale reversal of the shift toward automated methods of warfighting. While theoretically possible and perhaps even ideal, the expedience of such methods makes this reply of questionable utility. Are there other options?

Perhaps military institutions will compensate for the loss of moral skills in combat personnel by instead cultivating them in the software engineers responsible for programming automated systems to act ethically. But this does not answer the question of how, or through what new professional practices, software engineers could gain the needed moral expertise. Professional education would not be sufficient - the study of ethics textbooks, articles on just war theory or legal briefs on international laws of war do not by themselves enable skillful moral *action* or virtue—only repeated practice of the activities those books describe can produce the requisite capacities. Wargames or virtual-reality simulators might aim to engender in programmers and supervisors of automated systems the required habits and talents of moral discernment; but it is highly questionable whether simulations would carry the situational richness and moral gravity that produces genuine virtue.

Perhaps the best option is to restrict the deployment of automated methods of warfare to just those contexts in which human judgments are consistently and gravely inadequate *and* lead to morally intolerable error – while preserving robust opportunities elsewhere for, and expectations of, human soldiers to cultivate and exercise moral virtue in the conduct of war. This policy could be adopted alongside other uses of automated systems that create positive opportunities for moral

² Thanks to Don Howard for pointing this out in personal correspondence.

upskilling of military professionals. For example, rather than developing artificial moral intelligence that supplants human decision-making in the use of lethal force, artificial intelligence systems might instead be usefully deployed to provide soldiers with enhanced information about morally salient features of the battlefield, or to offer improved feedback concerning the alignment of soldiers' habits and decision patterns with norms of military honor, courage and restraint. However, such ethically constructive policies are unlikely to be pursued until and unless military leaders, educators, officers *and* the designers of automated systems jointly acknowledge the importance of preserving in military practice a developmental path for moral skills and virtues.

Acknowledgments

Thanks to Don Howard and John Sullins for helpful discussion of several of the key issues discussed in this paper.

REFERENCES

- Ackerman, Spencer. 2012. "Pentagon: A Human Will Always Decide When a Robot Kills You." *Wired*. Last modified November 26, 2012. <http://www.wired.com/dangerroom/2012/11/human-robot-kill/>.
- Aristotle. 1984. *The Complete Works of Aristotle: Revised Oxford Translation*. Edited by J. Barnes. Princeton: Princeton University Press.
- Augustine. 1994. *Political Writings*. Edited by Ernest Fortin and Douglas Kries. Indianapolis, IN: Hackett Publishing.
- Arkin, Ronald. 2009. *Governing Lethal Behavior in Autonomous Robots*. Boca Raton: CRC Press.
- Arkin, Ronald. 2010. The Case for Ethical Autonomy in Unmanned Systems. *Journal of Military Ethics* 9:4, 332-41.
- Asaro, Peter M. 2008. "How Just Could a Robot War Be?" In *Current Issues in Computing and Philosophy*. Edited by Philip Brey, Adam Briggles and Katinka Waelbers, Amsterdam: IOS Press. 50-64.
- Asaro, Peter M. 2009. "Modeling the Moral User." *IEEE Technology and Society Magazine* 28:1, 20-4.
- Axtell, Guy and Olson, Philip. 2012. "Recent Work in Applied Virtue Ethics." *American Philosophical Quarterly* 49:3, 183-203.
- Braverman, Harry. 1974. *Labor and Monopoly Capital: The Degradation of Work in the Twentieth Century*. New York: Monthly Review Press.
- Carter, Ashton B. 2012. "Department of Defense Directive: Autonomy in Weapons Systems." Last modified November 21, 2012. www.dtic.mil/whs/directives/corres/pdf300009p.pdf.
- French, Shannon E. 2003. *The Code of the Warrior: Exploring Warrior Values Past and Present*. New York: Rowman and Littlefield.
- Human Rights Watch. 2012. "Losing Humanity: The Case Against Killer Robots." Last modified November 19, 2012. <http://www.hrw.org/reports/2012/11/19/losing-humanity-0>.

- Krishnan, Armin. 2009. *Killer Robots: Legality and Ethicality of Autonomous Weapons*. Surrey, UK: Ashgate Publishing Limited.
- Lin, Patrick, Abney, Keith and Bekey, George A. 2012. *Robot Ethics: The Ethical and Social Implications of Robotics*. Cambridge, MA: MIT Press.
- Manders-Huits, Noemi. 2006. "Moral Responsibility and IT for Human Enhancement." *Proceedings of the 2006 ACM Symposium on Applied Computing*, 267-71.
- Olsthoorn, Peter. 2011. *Military Ethics and Virtues: An Interdisciplinary Approach for the 21st Century*. New York: Routledge.
- Roberts, John. 2010. "Art After Deskillling." *Historical Materialism* 18: 77-96.
- Reichberg, Gregory M., Henrik Syse and Endre Begby. 2006. *The Ethics of War: Classic and Contemporary Readings*. Oxford: Blackwell Publishing.
- Rinard, Ruth G. 1996. "Technology, Deskillling and Nurses: The Impact of the Technologically Changing Environment." *Advances in Nursing Science* 18:4, 60-9.
- Robinson, Paul. 2007. "Magnanimity and Integrity as Military Virtues." *Journal of Military Ethics* 6:4, 259-69.
- Royakkers, Lambèr & van Est, Rinie. 2010. "The Cubicle Warrior: The Marionette of Digitalized Warfare." *Ethics and Information Technology* 12:3: 289-96.
- Shakarian, Paulo, Devon Callahan, Jeffery Nielsen and Anthony N. Johnson. 2012. "Shaping Operations to Attack Robust Terror Networks." *Human Journal* 1: 15-25.
- Sharkey, Noel. 2010. "Saying 'No!' to Lethal Autonomous Targeting." *Journal of Military Ethics* 9:4, 369-83.
- Sharkey, Noel. 2012. "Automating Warfare: Lessons Learned from the Drones." *Journal of Law, Information and Science* 21:2.
- Singer, Peter W. 2009. *Wired for War: The Robotics Revolution and 21st Century Conflict*. New York: Penguin Group.
- Sparrow, R. (2007). Killer robots. *Journal of Applied Philosophy* 24:1, 62-77.
- Toner, James H. 2000. *Morals Under the Gun: The Cardinal Virtues, Military Ethics and American Society*. Lexington, KY: University Press of Kentucky.
- Townsend, Keith and Charles, Michael B. 2008. "Jarhead and Deskillling in the Military: Potential Implications for the Australian Labour Market." *Australian Bulletin of Labour* 34:1, 64-78.
- United States Air Force. 2009. *Unmanned Aerial Systems Flight Plan 2009-2047*. Last modified May 18, 2009: Washington, D.C. <http://www.govexec.com/pdfs/072309kpl.pdf>.
- Vallor, Shannon. 2013 (forthcoming). "Armed Robots and Military Virtue." In *Ethics of Information Warfare*. Edited by Luciano Floridi and Mariarosaria Taddeo, New York: Springer.