AUTOMATED WARFARE

George R. Lucas, Jr.*

In this Article, I review the military and security uses of robotics and "unmanned" or "uninhabited" (and sometimes "remotely piloted") vehicles in a number of relevant conflict environments that, in turn, raise issues of law and ethics that bear significantly on both foreign and domestic policy initiatives. My treatment applies to the use of autonomous unmanned platforms in combat and low-intensity international conflict, but also offers guidance for the increased domestic uses of both remotely controlled and fully autonomous unmanned aerial, maritime, and ground systems for immigration control, border surveillance, drug interdiction, and domestic law enforcement. I outline the emerging debate concerning "robot morality" and computational models of moral cognition and examine the implications of this debate for the future reliability, safety, and effectiveness of autonomous systems (whether weaponized or unarmed) that might come to be deployed in both domestic and international conflict situations. Likewise, I discuss attempts by the International Committee on Robot Arms Control (ICRAC) to outlaw or ban the use of autonomous systems that are lethally armed, as well an alternative proposal by the eminent Yale University ethicist, Wendell Wallach, to have lethally armed autonomous systems that might be capable of making targeting decisions independent of any human oversight specifically designated "mala in se" under international law. Following the approach of Marchant, et al., however, I summarize the lessons learned and the areas of provisional consensus reached thus far in this debate in the form of "soft-law" precepts that reflect emergent norms and a growing international consensus regarding the proper use and governance of such weapons.

* Professor of Ethics & Public Policy, Naval Postgraduate School.

I. CONCEPTUAL FOUNDATIONS OF ETHICS & LAW FOR UNMANNED SYSTEMS

The period from 2007 to 2013 witnessed an enormous outpouring of work devoted to the ethical and (far less frequently) the legal implications of military robotics. The inspiration for these studies stemmed from both the tremendous advances in the technologies themselves and the consequent dramatic increase in their roles in the conduct of military conflicts in many regions of the world.

These studies encompass Australian philosopher Robert Sparrow's inaugural essay, *Killer Robots*, and a subsequent, similarly titled book by Arman Krishnan.[1] A detailed and path-breaking survey of the ethical dilemmas posed by the increased use of such technologies prepared for the U.S. Office of Naval Research (ONR) by renowned computer scientist and roboticist, George Bekey, and his philosophy colleagues Patrick Lin and Keith Abney at California State Polytechnic University[2] heralded, in turn, the widely read and enormously influential treatment of the emerging ethical challenges and foreign policy implications of military robotics, *Wired for War*, by Brookings Institution senior fellow Peter W. Singer.[3]

The vast majority of these works focus on the ethical ramifications attendant upon the increased military uses of robotic technologies, reflecting the relevant lack of attention by legal scholars to the status of robotics in international law. The current status of domestic and international law governing robotics, however, together with a range of proposals for effective future governance of these technologies, was recently articulated by Arizona State University Law Professor, Gary Marchant, and several colleagues in the Consortium on Emerging Technologies, Military Operations, and National Security (CETMONS).[4] The legal and moral implications of military robotics constituted the main focus of a special issue of the *Journal of Military Ethics*[5] and of subsequent anthologies edited by Lin, Abney, and Bekey and, most recently,

---

1. ARMIN KRISHNAN, KILLER ROBOTS: LEGALITY AND ETHICALITY OF AUTONOMOUS WEAPONS (2009); Robert Sparrow, *Killer Robots*, 24 J. APPLIED PHIL. 62 (2007).

2. PATRICK LIN ET AL., AUTONOMOUS MILITARY ROBOTICS: RISK, ETHICS, AND DESIGN (2008), *available at* http://ethics.calpoly.edu/ONR_report.pdf.

3. P. W. SINGER, WIRED FOR WAR: THE ROBOTICS REVOLUTION AND CONFLICT IN THE TWENTY-FIRST CENTURY (2009).

4. Gary E. Marchant et al., *International Governance of Autonomous Military Robots*, 12 COLUM. SCI. & TECH. L. REV. 272 (2011). For more recent works, see Michael N. Schmitt, *Autonomous Weapon Systems and International Humanitarian Law: A Reply to the Critics*, HARV. NAT'L SEC. J. FEATURES 1 (Feb. 5, 2013), *available at* http://harvardnsj.org/wp-content/uploads/2013/02/Schmitt-Autonomous-Weapon-Systems-and-IHL-Final.pdf; and Jeffrey S. Thurnher, *The Law That Applies to Autonomous Weapon Systems*, 17 ASIL INSIGHTS 1 (Jan. 18, 2013), *available at* http://www.asil.org/sites/default/files/insight130118.pdf.

5. *New Warriors and New Weapons: Ethics & Emerging Military Technologies*, 9 J. MIL. ETHICS 1 (2010).

Bradley J. Strawser, along with a book-length treatment of aerial robotics ("drones") from an operational perspective by Colonel M. Shane Riza, USAF.[6]

It is worth pausing to reflect on what we have learned about the legal, moral, and policy implications of these trends as a result of these numerous and substantial efforts. First, while the technologies themselves are designed to operate in the domains of air, land, and sea, as well as in space, the majority of the discussion has centered on "unmanned" or remotely piloted, lethally armed aerial platforms (such as Predators and Reapers). That in turn stems from the highly effective use of these aerial platforms in surveillance operations, sometimes resulting in "targeted killing" of selected high-value adversaries by the United States and its allies. Indeed, it is often difficult to disentangle the discussions of aerial robotic technologies either from their controversial tactical deployment in such operations, or from the long-term strategic consequences of America's willingness to engage in such tactics. The tactical uses and strategic consequences of these policies involving unmanned systems, however, are quite distinct from the moral dilemmas posed by the vastly wider development and use of these systems themselves.

It is particularly unfortunate that the otherwise important policy discussions and moral debates surrounding the longstanding practice of targeted killing tends to obscure the fact that some of the most effective and justifiable uses of military robotics have been in unarmed ground operations, ranging from exploration of booby-trapped caves in Tora Bora, to battlefield rescue and casualty extraction, to dismantling IEDs or assisting in humanitarian relief operations.[7] Meanwhile, some of the most promising future developments in military

---

6. KILLING BY REMOTE CONTROL: THE ETHICS OF AN UNMANNED MILITARY (Bradley Jay Strawser, ed., 2013); M. SHANE RIZA, KILLING WITHOUT HEART: LIMITS ON ROBOTIC WARFARE IN AN AGE OF PERSISTENT CONFLICT (2013); ROBOT ETHICS: THE ETHICAL AND SOCIAL IMPLICATIONS OF ROBOTICS (Patrick Lin, Keith Abney & George A. Bekey eds., 2011); *see also* Heather M. Roff, *Killing in War: Responsibility, Liability and Lethal Autonomous Robots*, *in* ROUTLEDGE HANDBOOK OF ETHICS AND WAR: JUST WAR THEORY IN THE 20TH CENTURY 348-364 (Fritz Allhoff et al. eds., 2013); Joshua Foust, *A Liberal Case for Drones*, FOREIGN POLICY (May 14, 2013), http://www.foreignpolicy.com/articles/2013/05/14/a_liberal_case_for_drones.

7. A sense of the range of applications available can be found in many places, from the Unmanned Systems Roadmap of the U.S. Department of Defense, to the Unmanned Aeriel Systems online newsletter containing articles documenting the present and near-future use of unmanned systems for mapping coral reefs, monitoring the arctic environment, and apprehending drug trafficking. DEPARTMENT OF DEFENSE, UNMANNED SYSTEMS ROADMAP: 2007-2032 (2007); UAS VISION, http://www.uasvision.com/ (last visited Jan. 8, 2014). An upcoming small systems business expo includes feature articles on the uses of unmanned systems in humanitarian relief operations in Haiti, fighting forest fires, and agricultural crop spraying. SMALL UNMANNED SYSTEMS BUSINESS EXPO, http://susbexpo.com/ (last visited Jan. 8, 2014). A recent GAO report detailing FAA preparations for these uses also cites risks and abuses, ranging from unsafe operations of model aircraft operating too close to pedestrians at the University of Virginia, to apprehension of a domestic terrorist who was rigging a similar model system to carry plastic explosives in a planned attack on the White House. GOV'T ACCOUNTABILITY OFFICE, UNMANNED AIRCRAFT SYSTEMS (2012), *available at* http://www.gao.gov/assets/650/648348.pdf. Descriptions of uses of unmanned systems for

robotics will likely be realized in the maritime and underwater environment (in surface combat or anti-submarine warfare, for example), as well as when some of these systems return home from the warfront and are employed in a variety of domestic or regional security operations (border security, immigration control, drug and human trafficking, kidnapping, or disaster response and relief following hurricanes, floods earthquakes and massive wildfires) to which scant attention has thus far been paid (apart from implied threats to individual privacy).

During the Cold War, for example, it was often the case that submarines from rival superpowers engaged in intelligence, surveillance, and reconnaissance (ISR) missions in contested waters, keeping an eye on the adversary's movements and interests in a particular sector, and perhaps playing games of "cat and mouse" and even "chicken," to assess everything from noise detection to the maneuverability and other technical capabilities of different models of submarines as well as to test the effectiveness of coastal defenses. With the demise of the superpower rivalry and the Cold War, however, it has been some time since any naval force could routinely expend the resources necessary to continue such Tom Clancy-like, macho underwater scenarios. They are simply too risky and resource intensive.

Our strategic focus, meanwhile, has shifted from the Atlantic to the Pacific, and from Russia to China, as our treaty partners like Japan, South Korea, and the Philippines contend with one another and with the Chinese mainland for control of resource-rich areas of the South China Sea. Here a more typical scenario would involve an underwater ISR mission near the Diayu/Sinkaku islands, carried out by the United States in support of the interests of one of our principal allies, like Japan or South Korea. That operation today can be more efficiently and cost-effectively undertaken by deploying a single-manned vessel as an ISR command center, equipped with a variety of unmanned underwater vehicles (UUVs), each programmed to operate semi-autonomously in carrying out a series of task-oriented maneuvers in much the same way, and even following much the same command or decision-tree script that human commanders would have followed in an earlier era.

In an underwater "runtime environment," for example, robots behave, or are programmed to behave, with about the same degree of autonomy as a human commander of an individual manned ISR platform: that is, the operational or mission orders are for either type of vehicle to search, find, report, and either

---

multiple purposes in the marine environment can be found in the 2012 annual report of the multi-institutional Consortium on Unmanned Systems Education and Research (CRUSER). CONSORTIUM ON UNMANNED SYSTEMS EDUCATION AND RESEARCH, FY12 ANNUAL REPORT (2012), *available at* http://lgdata.s3-website-us-east-1.amazonaws.com/docs/1314/612480/CRUSER_AnnualReport_FY2012.pdf. An enormous archive of 3-dimensional visual prototypes in all relevant environments can be found online. *See* SAVAGE X3D EXAMPLES ARCHIVE, https://savage.nps.edu/ (last visited Jan. 8, 2014) (requires installation of an extension 3-D (X#D) plug-in, available on this site).

continue the mission or return to the command center (or to another specified rendezvous point).[8] This kind of mission can prove to be dull, dirty, routinely dangerous (for the manned platform), and certainly boring, until, that is, an adversary's submarine is observed carrying out exploratory mining surveys on the ocean floor. In a plausible "war-game" scenario, we might posit that the adversary then attempts to evade detection by fleeing into a prohibited marine sanctuary under the administrative control of yet another party to the dispute (e.g. the Philippines). The hypothetical semi-autonomous UUV would then face exactly the same legal and moral dilemma as would confront the human commander of a conventional manned submarine under otherwise-identical circumstances: namely, to continue the military mission of tracking the enemy or to refuse to violate international law and norms by hesitating to enter these prohibited, "no-go" waters. Standard operating procedures and "standing orders" defining the mission would require the human commander to "contact operational headquarters" for clearance, or else discontinue the mission. The UUV can relatively easily be programmed to do likewise, thereby incorporating the constraints of law and morality within the parameters of the rules of engagement defining this well-defined (and what I have elsewhere termed "highly scripted") mission.[9]

This seems a clear and relatively straightforward situation, for which we have unfortunate precedent in the human case.[10] The underlying question of this hypothetical case is whether (perhaps as a result of this precedent) to build a corresponding "rule of engagement" constraint for this mission that would override the tactical priorities of military necessity in deference to the requirements of law and good international relations. What seems equally clear at present is that we lack the technical capability to design a UUV with sufficient independent decision-making capacity to simulate the "swagger" of one of Tom Clancy's human commanders during the Cold-War era, and to decide *on its*

---

8. Donald P. Brutzman et al., *Run-Time Ethics Checking for Autonomous Unmanned Vehicles: Developing a Practical Approach*, *in* PROCEEDINGS OF THE 18TH INTERNATIONAL SYMPOSIUM ON UNMANNED UNTETHERED SUBMERSIBLE TECHNOLOGY (August 13, 2013), *available at* https://savage.nps.edu/AuvWorkbench/documentation/papers/UUST2013 PracticalRuntimeAUVEthics.pdf.

9. I define highly scripted security missions utilizing lethally armed autonomous systems in George R. Lucas, Jr., *Industrial Challenges of Military Robotics*, 10 J. MIL. ETHICS 274 (2011). A similar account of where and how to attain design success in ethical machine behavior can be found in Robert Sparrow, *Building a Better Warbot: Ethical Issues in the Design of Unmanned Systems for Military Applications*, 15 J. SCI. & ENGINEERING ETHICS 169 (2009). The appeal for human executive oversight provides for the possibility of overriding the legal prohibition, but renders the unmanned system in this instance merely "semi-autonomous" rather than fully autonomous, because the accountability for the decision rests with the human commander approving the action, not the unmanned platform.

10. An actual U.S. Navy mine-sweeper took such a detour with unfortunate and widely publicized results. *See U.S. Navy Ship Runs Aground in the Philippines*, CBSNEWS.COM (Jan. 17, 2013), http://www.cbsnews.com/8301-202_162-57564485/u.s-navy-ship-runs-aground-in-the-philippines.

own to override this legal constraint and venture into the "no-go" zone in search of the "bogie."

Happily, meeting the normal demands of law and morality does not require such a complicated, and likely infeasible, feat of engineering. Moreover, were we to design and program our unmanned undersea system to perform in the more or less straightforward manner otherwise described, we would have both duplicated the conventional behavior of a fully manned system, while simultaneously fulfilling what I have elsewhere defined as the "Arkin Test" or Arkin-constraint for robot morality: that is, we would have succeeded in designing, programming, and deploying a well-governed, reliable unmanned system that could almost certainly perform as well or even better (from a moral and legal perspective) than human beings under similar or identical circumstances.[11]

Even if I am correct in this claim, it may seem that I have cheated or "fudged" the boundary conditions to arrive at this result. Apart from whales, dolphins, and perhaps the intrepid underwater explorer, James Cameron, there are not a large number of civilian noncombatants wandering the undersea environment. Just as with Russian and American submarines during the Cold War, it is primarily adversaries (and some military allies) that one is likely to encounter there, lessening considerably the prospect of accidentally or unintentionally doing harm to innocents (collateral damage). On balance, that is to say, the undersea environment affords a relatively uncomplicated moral context within which to operate in comparison, say, to that of soldiers in a brigade combat team (or lethally armed robots) entering a local village in search of insurgents, let alone of a UAV making its own, wholly unsupervised targeting decisions and executing a full-blown "kill-chain" command unilaterally in mid-flight. It is the latter prospects that have raised the specter of uncontrolled and morally unaccountable "killer robots" run amok, and led ethicists like Wendell Wallach (Yale University) to propose having such systems declared "*mala in se*,"[12] while roboticists like Noel Sharkey and his colleagues in ICRAC to de-

11. This criterion—that robots comply as or more effectively with applicable constraints of the Law of Armed Conflict (LOAC) on their use of force and doing of harm than human combatants under similar circumstances—constitutes what I have termed the "Arkin Test" for robot "morality" (although that is likewise somewhat misleading, as the criterion pertains straightforwardly to compliance with international law, not with the exhibiting of moral judgment). In this sense, the test for "morality" (i.e., for the limited ability to comply with legal restrictions on the use of force) is similar to the "Turing Test" for machine intelligence: we have satisfied the demand when machine behavior is indistinguishable from (let alone better than) human behavior in any given context. I have outlined this test in several places. *See, e.g.*, George R. Lucas, Jr., *Postmodern War*, 9 J. MIL. ETHICS 289 (2010). For greater detail, see George R. Lucas, Jr., *Engineering, Ethics, and Industry: The Moral Challenges of Lethal Autonomy*, *in* KILLING BY REMOTE CONTROL 221 (Bradley Jay Strawser ed., 2013).

12. Wendell Wallach, *Terminating the Terminator: What to Do about Autonomous Weapons*, SCIENCE PROGRESS (Jan. 29, 2013), http://scienceprogress.org/2013/01/terminating-the-terminator-what-to-do-about-autonomous-weapons.

mand that the very development, let alone deployment of lethally armed autonomous systems be outlawed altogether.[13]

The UUV in the present example, however, is not lethally armed (at least, not yet), and it is only semi-autonomous (although, as described, it possesses essentially as much autonomy as a human operator or commander is authorized to exercise). This may seem to make consideration of moral dilemmas and legal challenges appear overly simplified. *But that is not a bad thing*. If the issue of military or security robots behaving ethically, let alone of their coming to exercise a computational analogue of human moral judgment, seems a complex and controversial matter as well as a formidable engineering challenge, then *why not start with something simpler*? What this example of the underwater runtime environment demonstrates is that we might feasibly design, build, program, test, and ultimately deploy reliable systems with only minimal "intelligence" and autonomy that nonetheless meet the demands of law and morality when operating in their defined battlespace.

II. DEVELOPING APPROPRIATE HYPOTHETICAL CASE STUDIES

By continuing to focus our work even more precisely on this somewhat simplified moral environment, we might learn lessons about how subsequently to approach the more complicated problems on the water's surface, on land, and in the air, as well as learning what to anticipate when we enter these areas.

For example, we can add a bit more complexity to our working underwater scenario. What if our UUV continues to tail the enemy and is fired upon? Do we arm it defensively against this possibility, and if so, does it return fire? Or does it take evasive action only? Or does it abort the mission and return to the host ship? Once again, in another era, human commanders, both on the surface and underwater, routinely confronted such decisions and were guided in making them by situational awareness, as well as by knowledge of both the conflict-specific and general (or "standing") rules of engagement (ROEs) that constrain and inform such activities. The key similarity is that *full autonomy to act without external restraint was seldom granted to the human commanders*. That is to say, despite the human commanders possessing a wide range of complex and varying capacities—intuitive judgment, leadership, moral scruples, conscience—that would be extremely difficult to replicate in machine behavior, the humans in fact usually simply acted in accordance with these Standing Orders. Generally, humans acted within a regulatory and governance framework rang-

13. *See* Noel Sharkey, *Saying 'No!' to Lethal Autonomous Targeting*, 9 J. MILITARY ETHICS 369 (2010). For an account of the work of ICRAC, see Nic Fleming, *Campaign Asks for International Treaty to Limit War Robots*, NEW SCIENTIST, (Sept. 30, 2009 3:41 PM), http://www.newscientist.com/article/dn17887-campaign-asks-for-international-treaty-to-limit-war-robots.html. *See also Berlin Statement,* INT'L COMMITTEE FOR ROBOT ARMS CONTROL, http://icrac.net/statements (last visited Jan. 8, 2014); *Mission Statement*, INT'L COMMITTEE FOR ROBOT ARMS CONTROL, http://icrac.net/who/ (last visited Jan. 8, 2014).

ing over a set of considerations, from the international law of the sea, to humanitarian law and a range of treaty obligations, all the way to specific rules of engagement designed to remove as much ambiguity and personal responsibility or accountability as possible from their actions (even if that is not exactly how they saw it, thought about it, or would have described it at the time). The human commanders had some situational autonomy, but not full (let alone unlimited) autonomy. Hence, their behaviors might be far easier to simulate with unmanned systems in the underwater military environment. Here, in this more clearly defined arena (as opposed to a range of other, more complex environments) we might hope at least to develop unmanned, semi-autonomous systems that could function at a level of reliability and safety equivalent to humans, even absent the advances in programming and governance technology that would otherwise be required to literally duplicate or replace the human presence or involvement in these scenarios.

If this vastly less presumptuous and far more attainable objective seems feasible in the case of underwater systems, might we then return to the even more complex land-based scenario of the brigade combat team engaged in a recon mission searching for insurgents in a nearby village? Like the underwater ISR mission, this scenario is for the most part "dull, dirty, and dangerous," as well as tedious and unrewarding, and on those grounds alone it is in principle amenable to the use of unmanned systems. The dull, dirty, and boring aspects cease, however, whenever something dramatic and unexpected occurs, like a small child bursting out of a house and running in fear across the path of an enemy insurgent who is about to fire his rifle or escape. Could an unmanned ground system conceivably cope with that foreseeable but unexpected circumstance, and others like it, with the degree of flexibility and sound judgment we demand of human combatants? Specifically, could it reliably and safely recognize the child as a noncombatant, and appropriately withhold fire (even if the legitimate target is thereby allowed to fire freely or escape capture)?

In the underwater case, we compared the degree of autonomy and sophistication required for the unmanned system to duplicate, not the full capacities of the human commander, but what we might term *the capacities specifically authorized for use* by that commander (which are considerably less than the full range available in the human case). Likewise, in the village scenario, we would compare the capacities required of a functioning and semi-autonomous ground system to those required and expected of, say, a young army private, recently recruited, trained, oriented to the ROEs applicable, and deployed in this conflict zone. Do we have a reasonable prospect of designing an unmanned ground system to function as well as, say, a relatively new army private? More generally, is there a reasonable prospect for designing and integrating unmanned ground systems into the "force mix" of brigade combat teams in order to work reliably alongside, or even replace, some of the army privates (if not the sergeant or lieutenant)?

There are many more instances, as well as variations of these types of unexpected encounters or unforeseen developments in the land-based case than in the underwater environment. The robot has some surprising advantages over the army private when it comes to its potential reaction time, especially with respect to the risk of harm it can tolerate in comparison to the human combatant. Where the unmanned system presently lags far behind the army private—and why the deployment of safe and reliable land-based systems may require a much longer period of development—is in what I described in the underwater instance as its degree of "situational awareness." Automatic target recognition or pattern recognition software employed in the unmanned system needs to be able interact quickly and reliably with hardware sensors to enable the system to distinguish between a child and an adult, between an insurgent and a shepherd, and between an AK-47 and a child's toy or a shepherd's crook. While this can in fact be done at present, such distinctions cannot yet be made at a fast operational pace with anything like the reliability of the army private, even when allowing for the occasional errors in the latter's situational awareness amid the "fog of war."

But what if one day soon such systems *could* reliably and consistently demonstrate the requisite degree of situational awareness? Or what if, in the meantime, we chose to develop and deploy land-based systems only in combat environments that resembled the very clearly structured underwater environment—i.e., characterized by very precise boundary conditions with little likelihood of the anomalous events and concomitant mistakes of judgment or perception that lead in turn to tragedy? The development of fully effective and reliable target and pattern recognition software and "friend or foe" systems may well lie in the reasonably near-term, foreseeable future. And meanwhile, robots have, in fact, been deployed in highly scripted contexts, such as serving as sentries and border guards in prohibited demilitarized or "no-go" zones. Robot sentries are currently used effectively in Israel as border guards in remote, strictly "entry-forbidden" areas, as well as by South Korea in the demilitarized zone with the North.[14] While primarily used for ISR at sea, the U.S. Navy is considering a plan to arm and deploy smaller versions of its "Fire Scouts"—unmanned and semi-autonomous helicopters—to provide force protection to ship convoys or aircraft carriers to shield them from insurgent attacks when approaching ports-of-call.[15] The development and increased use of unmanned systems in such environments could well lead to effective and relatively low-cost force multiplication with lower risk for combatants and noncombatants

14. *South Korea Deploys Robot Capable of Killing Intruders along Border with North*, THE TELEGRAPH (July 13, 2010), http://www.telegraph.co.uk/news/worldnews/asia/southkorea/7887217/South-Korea-deploys-robot-capable-of-killing-intruders-along-border-with-North.html.

15. Christopher P. Cavas, *U.S. Navy's New, Bigger Fire Scout to Fly this Fall*, DEFENSE NEWS (June 11, 2013 3:45 AM), http://www.defensenews.com/article/20130611/DEFREG02/306110009.

alike. If we can do these things, and if doing so works as well or better than our current practice, then perhaps we should pursue more fully our capabilities to do so. Moreover, if we determine that we can and should make these modifications, and if it is feasible to do so, ought we not to be trying as hard as we can to bring about these marked improvements in the conduct of armed conflict?

III. UNDERLYING PHILOSOPHICAL CONSIDERATIONS

This last line of reasoning describes another dimension of the Arkin Test: an unmanned platform fulfills the demands of law and morality (and may therefore be permissibly deployed) when it can be shown to comply with legal and moral requirements or constraints as well or better than a human under similar circumstances. Computer scientist Ronald C. Arkin, to whom we owe the benchmark, observes that this principle may also serve to generate *a technological obligation* to move forward with the development and use of robotic technology that would render war itself, and the conduct of armed hostilities, less destructive, risky, and indiscriminate.[16] The prospects for increasingly automated warfare turn critically on this moral claim.

Satisfying the Arkin Test, however, requires that we identify, analyze, and replicate the range of requisite behaviors in question sufficiently well enough to generate a reliable program script for an unmanned system that will emulate or duplicate the kind of human judgment and action called for in a given situation. In the underwater case, we can foresee, for example, that a human commander might be tempted to override the ROE prohibition against venturing into the international marine preserve because he or she "notices" something that our comparatively limited and literal-minded UUV would miss. But of course, the human commander might also override ROEs out of pride, ego, ambition, or just plain poor judgment. Likewise, in the village, the army private might exercise what partisans of folk psychology would label "intuitive judgment" in order to spare innocent lives, but might equally well instead react in just the opposite fashion—motivated perhaps by confusion and fear, racism, hatred, resentment, or mental disturbance—leading to an excessive and indiscriminate use of force resulting in the tragic injury or death of noncombatants.

While on the one hand we probably cannot analyze the intuition and situational awareness of the human into simplified Boolean logic, we have the compensating reassurance on the other hand that unmanned systems could not conceivably emulate any of the other undesirable human reactions described above. This is simply because UUVs or UGVs do not care, they have no interests, intentions, or self-regard, they harbor no ambitions or hatred, and they are utterly incapable of the "interiority" (to use another metaphor of folk psycholo-

16. Ronald C. Arkin, *The Case for Ethical Autonomy in Unmanned Systems*, 9 J. MIL. ETHICS 332, 347-56 (2010); *see also* RONALD C. ARKIN, GOVERNING LETHAL BEHAVIOR IN AUTONOMOUS ROBOTS (2009) (setting forth Arkin's definitive perspective on the topic).

gy) characteristic of self-consciousness. The German philosopher Martin Heidegger famously maintained that "Care [*Sorge*] is the Being of *Dasein*," a complex manner of depicting that such "interior" features—concern and compassion, as well as hatred and ambition, self-deception and self-regard—constitute the essence of being human.[17]

Heidegger's insight pertains especially well to the elusive folk term "conscience," that seems to integrate several of the components I have mentioned specifically, alongside others. I care about others than just myself, I have attachments of friendship, love and loyalty, and I sense that there are bonds and expectations connected to these that generate a range of duties and obligations to act in certain ways and avoid others. And, when tempted to override those constraints for the sake of immediate mission accomplishment (alternatively described more straightforwardly and less euphemistically as self interest, expediency, or personal gain), my acknowledgement of these other concerns causes cognitive dissonance, resulting in uncertainty about how to proceed. Because I "care" about these other matters and persons, I feel guilt, which may intervene to impede my attempts to override those concerns for the sake of expediency—or, failing to do so, the resulting "guilt" may function as a kind of biofeedback to improve my behavior in the future. This is quintessentially human. Our own, human caring is one (perhaps highly effective) manner of behavioral governance, in both an individual and social sense. We might just as well say that this trait of "caring," along with the constellation of intentional and emotional states surrounding it, constitutes the unique software package with which we have come from the factory pre-loaded, and that this particular software package works with our particular hardware effectively in most cases to generate responsible, intentional, accountable individual behavior within a wider social or cultural context. And whether we might additionally speculate that God, or evolution, or both, or something else entirely has conspired to design this arrangement, regardless it still remains functionally the case that emotions like guilt and concomitant phenomenal experiences like "conscience" function as behavioral modifiers. They are part of an elaborate and complex human feedback system that serves to modulate, constrain, and modify our individual behavior in a social context.

One strategy in robotics that takes account of the foregoing observations is, accordingly, to pursue the path of "strong artificial intelligence," in which human engineers and computer scientists seek to develop computational models

---

17. This famous line occurs in Chapter Six of Heidegger's great early work, MARTIN HEIDEGGER, BEING AND TIME 225 (John Macquarrie & Edward Robinson trans., Harper & Row 1962) (1929). This is likely not the proper venue for entering into an extensive analysis of what this may portend for human existence, save to say that all the subtle forms of experience Heidegger describes in his own, Husserlian-like account of time-consciousness in the human case clearly are *not* part of the "experience" of machines, including robots. Indeed, one account of the essence of such artifacts is that they don't *have* "experiences." And insofar as these experiences are (as both Heidegger and his teacher Edmund Husserl aver) constitutive of human being, robots are not at all like, nor likely to be like, human beings.

of human moral behavior by attempting to replicate emotions like guilt, as part of a larger effort to develop a fully functioning machine analogue of human moral conscience. This highly ambitious strategy might entail trying to replicate other features of the complex pallet of human mental states or human tactics associated with the "phenomenon" of morality (and immorality), such as ambition, self-regard, or the ability to engage in deceptive and misleading behavior.[18] The ambitious agenda of "strong AI" has been advocated by a number of scientists and engineers, from Marvin Minsky[19] to Arkin himself, as the most promising avenue to designing machine morality.[20] As an ambitious, long-range research project, such initiatives have much to commend them. It would be fascinating to analyze "guilt" as an effective form of system feedback and behavioral modification and self-correction or "machine learning." Likewise, it seems clear that some machine analogue of deceptive behavior might sometimes prove necessary in a robotic system designed for sophisticated battlefield extraction, or perhaps elder care (guiding the machine's actions toward reducing fear, bewilderment, and shock in the treatment of battlefield casualties or in the care of dementia patients). Do we need to reproduce exactly, or even approximately, these identical capacities in unmanned systems in order to make them reliable and safe and effective? From the previous (albeit simplified) scenarios, it would appear not.

Robots are not, nor could they be, *nor would we need or wish to them to be*, "human" in this sense. In contrast to Arkin's stated research agenda in particular, such ambitions do not seem to constitute a *logically necessary* preliminary step toward guaranteeing safety, reliability, and the most basic compliance on the part of present-day unmanned systems with prevailing legal and moral

18. Deception is a fascinating case in point, in that behavioral scientists have known for decades that deception is not unique to humans. It occurs in a variety of species, and is utilized as a tactic in pursuit of a variety of ends, from simple survival to the fulfilling of desires and ambitions (whether for food, sex, power, or to elicit the affection of others). Daniel Dennet, an eminent cognitive scientist and philosopher of the mind, is perhaps best known for his accounts over the years of the role of deception as an important dimension of intentionality in animal and human behavior. *See, e.g.*, DANIEL C. DENNETT, THE INTERNATIONAL STANCE (1987). In the human case, deception is a powerful tactic to operationalize strategies with a variety of objectives that might be defined as "successful mission outcome." For example calming a terrified wounded combatant with reassuring words so that he may be brought back from the front to a military hospital for treatment, or, in quite a different sense building some sort of "Potemkin Village" (such as General George Patton's "First United States Army Group" (FUSAG) during World War II) so that the enemy might be deceived regarding troop strength, placement, or tactical intentions.

19. MARVIN MINSKY, THE EMOTION MACHINE: COMMONSENSE THINKING ABOUT ARTIFICIAL INTELLIGENCE, AND THE FUTURE OF THE HUMAN MIND (2006).

20. Ronald C. Arkin & Patrick Ulam, *An Ethical Adaptor: Behavioral Modification Derived from Moral Emotions*, *in* IEEE INTERNATIONAL SYMPOSIUM ON COMPUTATIONAL INTELLIGENCE IN ROBOTICS AND AUTOMATION 381 (2009); Ronald C. Arkin, Patrick Ulam & Alan R. Wagner, *Moral Decision Making in Autonomous Systems: Enforcement, Moral Emotions, Dignity, Trust, and Deception*, 100 PROC. IEEE 571 (2012).

norms of behavior in combat or security operations.[21] Once again, as the foregoing scenarios demonstrate, satisfactory results can be obtained employing existing software, hardware, and programming languages. Those modest successes in relatively uncomplicated situations might well lead researchers and end users toward further, fully feasible, and largely conventional engineering solutions to the more complex behavioral challenges in unmanned systems governance.

We might usefully compare this alternative research agenda with the history of aviation, in which initially ambitious and largely unsuccessful attempts to model animal behavior directly gave way to more modest and ultimately successful attempts to attain flight through the use of conventional materials and system designs that bore little resemblance to the actual behavior of birds or insects. Indeed, it is only within the past few years that engineering and material sciences (including miniaturization of powerful computing) have enabled engineers to actually model and duplicate the flight behaviors of birds and insects directly. Had we, however, demanded adherence to this strict narrow-minded correspondence of principle at the dawn of the aviation age (rather than rightly abandoning it at the time as unattainable), we would likely still be awaiting our first successful flight, and, through this lack of creative imagination, have missed out on an entire intervening century of exciting aviation developments, uses, and benefits that bear only the most rudimentary resemblance to animal behavior.

IV. MORAL AND LEGAL IMPLICATIONS OF A LESS COMPLEX RESEARCH AGENDA

The foregoing observations and insights are essential to addressing the host of controversies that have otherwise arisen in the field of military robotics in particular. Indeed, from the emergence and increasing use of unmanned or "remotely piloted" vehicles to the advent of cyber war and conflict, the development of new and exotic military technologies has provoked fierce and divisive public debate regarding the ethical challenges posed by such technologies. Peter Singer and Noel Sharkey, in particular, have focused their criticisms especially upon Ronald Arkin, whom they accuse of offering false and technologically unattainable promises as justification for his advocacy of greater use of military robots in combat.

Their argument is, in effect, in line with the foregoing observation regarding uniquely human capacities, traits, and behaviors that seem constitutive of

21. This observation is *not meant* to gainsay either the significance or utility of such research. Rather, I merely mean to indicate that Arkin's research agenda does not represent the only way forward on the question of "robot morality." His is an example of what one might term a "visionary" approach, as opposed, frankly, to a more pedestrian (or "workmanlike") approach making use of available technological capacity in the present to address the pressing questions of law and ethics with respect to unmanned systems.

moral reasoning and effective moral judgment and decisionmaking. The computational machine analogues of these human behaviors and mental states have not yet been attained, and will not likely be so in the near future, if ever. Indeed, both Singer and Sharkey think the promises of "strong" artificial intelligence researchers in these areas, in particular, are hyperbole and pure science fiction.[22] As a result, they claim there is no warrant for continuing to develop autonomous, lethally armed military hardware, and that it is deceptive and cynical to offer such vague promises as grounds for continuing to support and fund such research. A machine without these requisite guidance systems, lethally armed, and otherwise autonomous (self-guided, independent of ongoing human executive oversight, and capable of making and executing "kill-chain" targeting decisions unilaterally) would invariably prove dangerously unreliable, and the very pursuit of such research is reckless, if not criminally negligent.[23] Sharkey and Singer go so far as to claim, as a result, that attempts to design such systems, let alone deploy them, should be outlawed altogether. As mentioned in passing above, Wendell Wallach has proposed a somewhat different strategy that would designate the use of such systems wholly apart from any form of human supervision, executive oversight, or accountability as off limits. In this way of thinking, if the use of *wholly unsupervised* autonomous systems is thought to be inappropriate, Wallach argues, one could place a legal limit on the use of lethally armed versions of unmanned systems by having them designated *mala in se* under international law. This legal designation would have the effect of grouping the use of lethally armed autonomous systems alongside the use of chemical and biological weapons of mass destruction, as serving no legitimate military purpose whatever. Armed unmanned systems could circumvent this legal restriction only if they remained fully under human control, with accountability for targeting decisions (including errors and any resulting collateral damage) ascribed solely to human agents. Apart from the sensitive legal terminology entailed in this proposal, it would otherwise have consequences largely indistinguishable from the current ban on the development and use of unsupervised lethally armed autonomous systems established in the Department of Defense directive of November 2012. The designation would have the additional effect of making the current ban permanent rather than temporary.[24]

22. Noel Sharkey, *The Automation and Proliferation of Military Drones and the Protection of Civilians*, 3 L. INNOVATION & TECH. 236 (2011); Noel Sharkey, *March of the Killer Robots*, LONDON TELEGRAPH (June 15, 2009), http://www.telegraph.co.uk/science/science-news/5543603/March-of-the-killer-robots.html.

23. *See* SINGER, *supra* note 3 (commenting on the bad faith and deceptive nature of moral reassurances from engineers like Arkin who are engaged in promoting the values and virtues of autonomous systems); *see also* Peter W. Singer, *The Ethics of Killer Apps: Why Is it So Hard to talk about Morality When It Comes to New Military Technology?* 9 J. MIL. ETHICS 299 (2010).

24. Wallach, *supra* note 12. The legal sensitivity stems from the association of lethally armed autonomous systems with other means and methods of warfare that are "evil in themselves" (*mala in se*), such as rape. It is less clear or convincing, however, that the reasons

Such ongoing and intractable controversies may point toward an altogether different problem. These intractable moral dilemmas, ironically, may be taken as evidence that the language of morality and ethics is serving us poorly in this context, further confusing, rather than helpfully clarifying or enlightening us on how best to cope with the continuing development and deployment of seemingly exotic new military technologies. To the complaint of opponents of military uses of robotics that such uses are immoral and are, or ought to be, declared illegal, proponents respond by attempting to promise, at least, that their creations "will be able to behave as or more ethically than humans." In fact, however, opposing parties involved in these discussions harbor distinctive and incompatible—and sometimes conceptually confused and unclear—notions of what "ethics" entails. From individual and culturally determined intuitions regarding right conduct through the achievement of beneficial outcomes all the way to equating ethics merely with legal compliance, this conceptual confusion results in frequent and virtually hopeless equivocation. Moreover, many scientists and engineers (not to mention military personnel) tend to view the wider public's concern with "ethics" as misplaced, and regard the invocation of "ethics" in these contexts as little more than a pretext for technologically and scientifically illiterate, fear-mongering, nay-saying luddites who simply wish to impede the progress of science and technology.

But why should we insist on invoking fear and mistrust, and posing allegedly "moral" objections to the development and use of unmanned systems, instead of defining clear engineering design specifications and operational outcomes that incorporate the main ethical concerns? Why not simply require that engineers and the military either design, build, and operate their unmanned systems to these exacting standards, if they are able, or else desist from manufacturing or deploying such systems until they succeed in satisfying these engineering specifications? Why engage in a science-fiction debate over the future prospects for artificial machine intelligence that would incorporate analogues of human moral cognition, when (as I have demonstrated above in the case of UUVs) what is required is far more feasible and less exotic: namely, machines that function reliably, safely, and fully in compliance with applicable international laws—such as the law of armed conflict (LOAC), when operating in wartime? And why insist that the development and use of such systems would constitute a "game changer" that ushers in a new mode of unrestricted warfare,

adduced for this classification would prove compelling, both because the analogy between autonomous systems and other familiar examples of means *mala in se* (rape, in particular) do not appear obvious, while the author's argument still seems to rest in part on the objection I have tried in this Article to discredit: namely, that machines cannot be held accountable for their actions. Biological and chemical weapons of mass destruction, in addition, are not so designed on account of the design or type of the weapon itself, but because the use of such weapons is thought to cause unnecessary injury and superfluous suffering. It is hard to see how this could possibly be the case with lethally armed, autonomous "drones," where the death or injury from a missile is presumably identical to that experienced from the same missile fired from a manned aircraft.

in which all the known laws and moral principles of armed conflict are rendered obsolete, when what is required is merely appropriate analogical reasoning to determine how the known constraints on conventional armed conflict might be usefully extrapolated to provide effective governance for these novel conditions?

The prospects for machine models of moral cognition, as we have discovered to this point, constitute a fascinating but as yet futuristic and highly speculative enterprise. The goal of developing working computational models of reasoning, including moral reasoning, may not prove altogether impossible, but the effort required will surely be formidable. "Morality" and moral deliberation for the present (as critics of military robotics contend) remain firmly in the domain of human experience, and likely will remain there for the foreseeable future. In any event, discussions of ethics and morality pertaining to unmanned systems at present are largely irrelevant. We neither want nor need our unmanned systems to "be ethical," let alone "more ethical" or "more humane" than human agents. We merely need them to be safe and reliable, to fulfill their programmable purposes without error or accident, and to have that programming designed to conform to relevant international law (LOAC) and specific rules of engagement (ROEs). With regard to legal compliance, that is to say, machines should be able to pass the Arkin Test: autonomous unmanned systems must be demonstrably capable of meeting or exceeding behavioral benchmarks set by human agents performing similar tasks under similar circumstances, as we have thus far shown them quite capable in the case of UUVs.

On the other hand, proposals at this juncture simply to "outlaw" research, development, design, and manufacture of autonomous weapons systems seem at once premature, ill-timed, ill-informed—classic examples of "poor governance." Such proposals do not reflect the concerns of the majority of stakeholders who would be affected; they misstate, and would attempt to over-regulate relevant behaviors.[25] Ultimately, such regulatory statutes would prove unacceptable to and unenforceable against many of the relevant parties (especially among nations or organizations with little current regard for international law),

25. In addition to proposals to outlaw armed or autonomous military robotic systems by ICRAC itself. *See* HUMAN RIGHTS WATCH, LOSING HUMANITY: THE CASE AGAINST KILLER ROBOTS (2012), *available at* http://www.hrw.org/sites/default/files/reports/arms1112ForUpload_0.pdf. While unquestionably well-intentioned, the report is often poorly or incompletely informed regarding technical details, and highly misleading in many of its observations. Furthermore, its proposal for states to collaborate in banning the further development and use of such technologies would not only prove unenforceable, but likely would impede other kinds of developments in robotics (such as the use of autonomous systems during natural disasters and humanitarian crises) that the authors themselves would not mean to prohibit. It is in such senses that these sorts of proposals represent "poor governance." For a well-informed and clearly argued rejoinder to this report, critiquing the legal efforts to ban autonomous systems, see KENNETH ANDERSON & MATTHEW WAXMAN, LAW AND ETHICS FOR AUTONOMOUS WEAPON SYSTEMS: WHY A BAN WON'T WORK, AND WHY THE LAWS OF WAR CAN (2013).

and would thus serve merely to undermine respect for the rule of law in international relations. Machines themselves (lacking the requisite features of folk psychology, such as beliefs, intentions, and desires) by definition cannot themselves commit war crimes, nor could a machine itself be held meaningfully accountable for its actions under the law. Instead, a regulatory and criminal regime, respecting relative legal jurisdictions, already exists that holds human in-individuals and organizations who might engage in reckless and/or criminally negligent behavior in the design, manufacture, and end use of unmanned systems of any sort fully accountable for their behavior and its consequences.

V. Ethical Principles for Unmanned Systems Research and Deployment Policy

Rather than engaging in a headstrong and ill-informed rush to proposing unenforceable treaties or legislating more ineffectual "bright line" statutes of black-letter international law, the proper course of action would instead be to invite those engaged in the development and use of such technologies, in the course of their activities, to reflect upon and observe what appear to them to be the boundaries of acceptable and unacceptable conduct, and to codify these by consensus and agreement as the principles of what would be termed "best practices" in their fields. Indeed, what international relations and policy experts sometimes term "emergent norms" regarding ethics, legal jurisdiction, compliance, and appropriate degrees of consent and accountability for all the stakeholders already have been implicitly established. What is urgently needed at this juncture is a clear summary of the results of these ongoing discussions and debates that would, in turn, help to codify what we seem to have proposed or agreed upon in these matters, and what requires still further deliberation and attention.

In the case of the debate over autonomous systems, the unrecognized and neglected areas of mutual consensus during the past several years of contentious debate regarding unmanned systems and automated warfare may be usefully summarized in the following Precepts, which define good or best practices, and delineate the limits of acceptable versus unacceptable practice. The point of the exercise is not to presume or preempt proper legislative authority, but instead to focus future discussions upon whether such Precepts are correctly stated (and if not, to modify them accordingly), examine the degree to which these areas of presumed consensus are in fact widely held, and, finally, to identify areas of omission that must still be addressed. In that spirit, I offer these Precepts as a working summary to date of the emerging consensus on the use of autonomous weapons systems, and propose that a summit of practitioners and international stakeholders (in a procedure analogous to the Asilomar Convention on Genetic Research, or the Pugwash Conventions on Nuclear Arms)[26] be convened at the earliest possible date to continue to revise, amend, and agree to abide by these precepts governing the development and use of unmanned systems.

I. *The Principle of Mission Legality*

A military mission that has been deemed legally permissible and morally justifiable on all other relevant grounds does not lose this status solely on the basis of a modification or change in the technological means used to carry it out (i.e., by removing the pilot from the cockpit of the airframe, or replacing a submarine crew and commander with demonstrably reliable software), unless the technology in question represents or employs weapons or methods already specifically proscribed under existing international Weapons Conventions or that are in violation of the prohibitions in international humanitarian law against means or methods that inflict superfluous injury or unnecessary suffering (or the technology is otherwise judged to constitute means of warfighting that are *mala in se*).

II. *The Principle of Unnecessary Risk*[27]

Within the context of an otherwise lawful and morally justified international armed conflict or domestic security operation, we owe the combatant or domestic security agent every possible minimization of risk we can provide them in the course of carrying out their otherwise legally permissible and morally justifiable missions. We owe third parties and bystanders (civilian noncombatants) caught up in the conflict every protection we can afford them through the use of ever-improved means of conflict that lessen the chance of inadvertent injury, death, or damage to their property and means of livelihood.

[*Comment*: This precept combines the original insight of Strawer's principle of unnecessary risk, with Arkin's sense that military technologists should

26. A landmark conference on recombinant DNA research organized by geneticist Paul Berg at Asilomar, California in 1975 resulting in a moratorium on genetics research that might inadvertently generate serious harm in the absence of a full scientific understanding of the risks involved in such research. Earlier, nuclear scientists who desired to play a strong role in the policy implications of the proliferation and possible use of nuclear weapons of mass destruction in the 1950s and 1960s organized a series of working conferences, beginning at a site in the small town of Pugwash, Nova Scotia (Canada) to promote international efforts to limit the prospects of all-out nuclear war. Philosophy Professor Don Howard, Director of the Reilly Center for Science, Technology and Values at Notre Dame University, has proposed something like this process as an approach to ameliorating cyber conflict. I agree, and think it even more appropriate to the current status of the debate over ethics, law, and unmanned systems.

27. First formulated by Bradley J. Strawser. *See* Bradley J. Strawser, *Moral Predators: the Duty to Employ Uninhabited Aerial Vehicles*, 9 J. MIL. ETHICS 342 (2010).

feel obliged to exercise their expertise to lessen the cruel and anti-humanitarian aspects of armed conflict.]

III. *The Principle of the Moral Asymmetry of Adversaries*

No obligation of "fairness" or symmetry of technological advantage is owed to opponents or adversaries whenever the latter are unmistakably engaged in unjust or illegal use of force, whether during the commission of domestic crimes, or when involved in international criminal conspiracies (terrorism).

[*Comment*: It is sometimes mistakenly asserted that in international war and armed conflict, at least, there is some such obligation, and hence that one moral objection to the use of unmanned systems is that "the other side doesn't have them." Technological asymmetry is not a new phenomenon, however, but rather an enduring feature of armed conflict. No such constraint is imposed on domestic law enforcement engaged in armed conflict with, for example, international drug cartels. Likewise, no such obligation of symmetry is owed to international adversaries when they are engaged in similar criminal activities: e.g., violation of domestic legal statutes within the borders of a sovereign state, defiance of duly elected legal authorities, indiscriminate targeting of civilians and destruction of property, kidnapping, torture, execution, mutilation of prisoners, etc.][28]

IV. *The Principle of Greatest Proportional Compliance*

In the pursuit of a legally permissible and morally justifiable military (or security) mission, agents are obligated to use the means or methods available that promise the closest compliance with the international laws of armed conflict (LOAC) and applicable rules of engagement (ROEs), such as non-combatant distinction (discrimination) and the economy of force (proportionality).

[*Comment*: This is another implication of Arkin's assertion of an obligation to use unmanned systems whenever they might result in greater compliance with international law, and in the lessening of human suffering in war. In this

28. Note that this is not an explicit rejection of the doctrine of the "Moral Equality of Combatants," an essential element in what Michael Walzer defines as the "War Convention." MICHAEL WALZER, JUST AND UNJUST WARS 44 (2d ed. 1977). Rather, it is a repudiation of a misplaced notion of "fairness in combat," according to which it would be unfair for one side in a conflict to possess or use weapons or military technologies that afforded them undue advantage. This is sometimes cited in public as an objection to the use of "drones" in warfare. It seems to equate war with a sporting competition, after medieval jousting fashion, and, upon examination, is not only patently ridiculous, but contradicted in most actual armed conflicts of the past, where maneuvering for "technological superiority" was a key element in success. In any case, no such argument is made concerning legitimate domestic security operations, as noted above, and does not obtain either within the realm of wars of "law enforcement" or humanitarian intervention.

case, the implication is that nations involved in armed conflict must use the least destructive means available (whether these be robots, cyber weapons, precision-guided munitions, or "non-lethal" weapons) in pursuit of military objectives that are otherwise deemed to be morally justified and legally permissible.]

V. *The Arkin Test*

In keeping with Precept IV, an artifact (such as an autonomous unmanned system) satisfies the requirements of international law and morality pertaining to armed conflict or law enforcement, and may therefore be lawfully used alongside, or substituted for, human agents whenever the artifact can be shown to comply with the relevant laws and ROEs as (or even more) reliably and consistently as human agents under similar circumstances.

[*Comment*: Moreover, from application of Precepts II and IV above, the use of such an artifact is not merely legally permissible, but *morally required*, whenever its performance promises both reduced risk to human agents and enhanced compliance with LOAC and ROEs.]

VI. *The Principle of Non-Delegation of Authority and Accountability*

The decision to attack an enemy (whether combatants or other targets) with lethal force may not be delegated solely to an unmanned system in the absence of human oversight, nor may eventual accountability for carrying out such an attack be wholly abrogated by human operators otherwise normally included in the "kill chain."

[*Comment*: This Precept is indebted to the work of philosopher Robert Asaro of the New School (New York), co-founder of the International Committee for Robot Arms Control (ICRAC). It also brings professional canons of best practice in line with the requirements of the U.S. Department of Defense guidance on future unmanned systems, stating that autonomous unmanned systems shall not be authorized to make unilateral, unsupervised targeting decisions.][29]

VII. *The Principle of Due Care*

All research and development, design, and manufacture of artifacts (such as lethally armed and/or autonomous unmanned systems) ultimately intended for use alongside or in place of human agents engaged in legally permissible and morally justifiable armed conflict or domestic security operations must rigorously comply with Precepts I-V, above. All Research & Development (R&D), design, and manufacture of unmanned systems undertaken with full knowledge of, and in good faith compliance with, the above Precepts (such good faith at minimum to encompass rigorous testing to ensure safe and relia-

29. DEPARTMENT OF DEFENSE, *supra* note 7.

ble operation under the terms of these precepts) shall be understood as legally permissible and morally justifiable.

VIII. *The Principle of Product Liability*

Mistakes, errors, or malfunctions that nonetheless might reasonably and randomly be expected to occur, despite the full and good faith exercise of due care as defined in Precept VII above, shall be accountable under applicable international and/or domestic product liability law, including full and fair financial and other compensation or restitution for wrongful injury, death, or destruction of property.

[*Comment*: this practice mirrors current international norms as practiced by minimally rights-respecting States in the case of conventional armed conflict. When combatants accidentally or unintentionally bring about the injury or death of non-combatants, the responsible state acknowledges blame and offers apology and financial restitution to the victims or to their survivors in the aftermath of an investigation into the wrongful death or injury, which is undertaken to determine any additional criminal liability.]

IX. *The Principle of Criminal Negligence*

By contrast, R&D, design, or manufacture of systems undertaken through culpable ignorance, or in deliberate or willful disregard of these precepts (to include failure to perform, or attempts to falsify the results of, tests regarding safety, reliability of operation, and compliance with applicable law and ROEs, especially in the aftermath of malfunctions as noted above), shall be subject to designation as "war crimes" under international law, and/or as reckless endangerment or criminally negligent behavior under the terms of applicable international and/or domestic law. Individual parties to such negligence shall be punished to the full extent of the law, to include trial and conviction in the International Criminal Court for the willful commission of war crimes, and/or civil and criminal prosecution within the appropriate domestic jurisdiction for reckless endangerment or criminal negligence. In domestic jurisdictions providing for capital punishment upon conviction for the occurrence of such mishaps within that jurisdiction, such punishment shall be deemed an appropriate form of accountability under the Precepts above.

[*Comment*: This Precept incorporates the concerns, and addresses the objectives of critics of military robotics pertaining to wrongful injury, death, or destruction of property by unmanned systems in which a human combatant, under similar conditions, could and would be held criminally liable for the commission of war crimes. The precept allows imposition of the death penalty for such offenses when guilt is ascertained within legal jurisdictions permitting capital punishment.]

X. *Benchmarking*

Testing for safety and reliability of operation under the relevant precepts above shall require advance determination of relevant quantitative benchmarks for human performance under the conditions of anticipated use, and shall require any artifact produced or manufactured to meet or exceed these benchmarks.

[*Comment*: This operationalizes the otherwise vague concept of "the behavior of human beings under similar circumstances," requiring that this be ascertained and sufficiently well-defined to guide the evaluation and assessment of the requisite performance of unmanned systems proposed for use in armed conflict.]

XI. *Orientation and Legal Compliance*

All individuals and organizations (including military services, industries, and research laboratories) engaged in R&D, design, manufacture, acquisition, or use of unmanned systems for military purposes shall be required to attend an orientation and legal compliance seminar of not less than eight hours on these Precepts, and upon conclusion, to receive, sign, and duly file with appropriate authorities a signed copy of these Precepts as a precondition of their continued work. Failure to comply shall render such individuals liable under the principle of criminal liability (Precept IX) above for any phase of their work, including, but not limited to, accidents or malfunctions resulting in injury, death, or destruction of property.

Government and military agencies involved in contracting for the design and acquisition of such systems shall likewise require and sponsor this orientation seminar and facilitate the deposit of the required signed precept form by any contractors or contracting organizations receiving federal financial support for their activities. Federal acquisitions and procurement officials shall also receive this training, and shall be obligated to include the relevant safety/reliability benchmarks of human performance along with other technical design specifications established in RFPs or federal contracts.

[*Comment*: One frequently raised objection to the concept of "soft law" and governance according to emergent norms questions the degree of sanction and normativity attached to the enforcement of these norms. Inasmuch as the Precepts themselves define a region of criminal behavior, and establish the bounds of jurisdiction pertaining to criminal activity in these cases, this final Precept ensures that affected stakeholders deemed to be bound by these Precepts are fully cognizant of their content, interpretation, and prospective normative force.]

CONCLUSION

Law and moral discourse, famously, always lag behind technological innovations (especially, if not exclusively, in warfare) and their transformative impact on the cultures in which they arise. That does not mean that law and morality are irrelevant and must be cast aside, nor does it require that ethics always be portrayed as an impediment or obstacle to technological development. Rather it demands, as such developments always have, that human agents employ appropriate ingenuity in the framing of suitable metaphors, the drawing of the most appropriate analogies, and reasoning by extrapolation from the known to the unknown in the continuing quest to order and organize the perplexing opportunities and risks that innovation and change otherwise invariably pose.

Accordingly, my intent in offering these Precepts is to suggest areas of consensus and agreement discerned among contending practitioners and stakeholders in this debate about the increasing military reliance upon unmanned systems, and to suggest the norms emerging from this debate that might serve to guide (if not strictly govern) the behavior of states, militaries, and those involved in the development, testing and manufacture of present and future unmanned systems. I likewise contend that discussion of the meaning, application, and refinement of these precepts as "soft-law" guidelines for proper use of unmanned systems would be substantially more efficacious than further moral "hand-wringing" over their potential risks, let alone a rush to legislation that would have both unenforceable and unintended harmful consequences. I have personally advocated instead for a summit of practitioners and stakeholders to develop, improve, and disseminate guidance (such as these Precepts) emerging from practice and from the debate on the limits of acceptable practice in the contrasting views of current stakeholders. This would mirror the best practices of geneticists in the 1970s and 80s (the Asilomar Moratorium) and of nuclear scientists in the 1950s and 1960s (the Pugwash Conventions) in designing and disseminating what proved to be highly effective precautions for their respective investigations.

I have attempted to demonstrate that there is far more consensus than we have been able to discern among adversarial parties arguing about ethics and law in such matters than we have heretofore been able to discern. That emerging consensus, in turn, points toward a more productive regime of governance and regulation to assure against the risk of unintended harm and consequences than do rival attempts at legal regulation or moral condemnation.