

DRAFT: DO NOT DISTRIBUTE OR CITE WITHOUT PERMISSION

Jus nascendi, Robotic Weapons & the Martens Clause

Peter M. Asaro

The New School, asaro@newschool.edu

In this paper I want to consider the motivations and sources for new law, *jus nascendi*. In particular, I want to consider the motivations and sources for new International Humanitarian Law (IHL) concerning the development and use of robotic autonomous weapons. Among the motivations for *jus nascendi* in IHL, the development of new military technologies is one of the most pressing. The legality and potential regulation of robotic and autonomous weapons, alongside and sometimes in conjunction with cyberwarfare, have raised serious questions and a growing debate in the international community. While we are considering the various reasons and means for regulating autonomous weapons, I believe it is important to consider where new international law emanates from, particularly in response to new technologies.

The law itself is often viewed as largely stable and unchanging, though in fact it is dynamic in various ways. While interpretations and jurisprudence may evolve over time, the fundamentals of the law itself appear rather more stable. Of course, within nations new legislation is regularly adopted, but this is largely a matter of state policy and not an exact correlate to the introduction of new custom and treaty law in the international context. But even in international law, there is new and emergent customary and treaty law, and the dynamics of the practice and application of law leads to emergent jurisprudence and transformations in norms. In IHL, new technological developments are a key factor in these dynamics.

I have addressed some of the specific reasons for prohibiting the use of autonomous weapons elsewhere.¹ In this chapter I want to consider the legal framework and means by which new law could come into place for new robotic technologies, and some of the philosophical issues which arise in this process. Indeed, a great deal of energy has been spent trying to understand how, when and why new law might be necessary for autonomous weapons. I believe a clearer understanding of the general situation will advance not only the debate regarding autonomous weapons, but might also inform discussions on regulating other robotic technologies, a frequent issue of concern at the WeRobot conferences on robot law.

The concept of *jus nascendi* is of particular interest in the context of developing technologies and robot law in particular. The emergence of new technologies can and does challenge many of our existing assumptions and traditional interpretations of the law. The potential of robots to become agents in the world challenges long held assumptions built into the law that only humans can act

¹Peter Asaro, “On Banning Autonomous Lethal Systems: Human Rights, Automation and the Dehumanizing of Lethal Decision-making,” *International Review of the Red Cross*, 94 (886), Summer (2012), pp. 687-709. Downloaded from:

<https://www.icrc.org/eng/resources/documents/article/review-2012/irrc-886-asaro.htm>

as agents.² Robots thus present many cases and examples in various domains of law for which we must reconsider and refine our view of the law and its applicability. In the most extreme cases, this may mean revising the law itself, or writing law anew. This is most obvious in the laws designed to regulate new industries which might produce robotic products for applications from robot surgery, nursing and personal assistants, to robotic self-driving cars and surveillance drones.

Apart from these new consumer industries, the area in which the question of new law and new regulation is most pressing, and the robotic application domain where vast amounts of research and development is currently being directed, is in the development of robotic weapons. While these technologies could find use among civilian police forces and individuals, the primary aim of current research and development by states is towards sophisticated military weapons with increasing degrees of automation and information processing. It is thus in the context of IHL, which governs the use of weapons in international armed conflicts, that I will consider how the emergence of robotic weapons forces us to consider new norms and new law.

One particularly challenging aspect of new technologies is their potential for disruption. New technologies by their nature disrupt practices and behaviors. While the computerization of offices during the second half of the twentieth century may not have fundamentally changed the business that those offices conducted (though it arguably did), it did radically transform the day-to-day activities and practices of the people who conducted that business.³ Had computerization not done so, it would likely have failed to achieve the gains in productivity that the introduction of such technologies sought. The same can be said of the automation of factories in the century that proceeded computerization,⁴ leading to the introduction of the first numerically controlled machines and industrial robots.⁵

It is the increased automation and computerization of warfare, its weapons and the use of violent

²Of course the law recognizes human organizations and institutions—such as corporations, trusts, non-profits, and states themselves—as legal entities with certain legal responsibilities. However, the law still largely views these organizations and institutions through the agency of their employees, trustees, officers and executives. This is to say that the law still largely assumes that individual humans are making decisions and taking actions, even when they are acting in official capacities.

³Shoshana Zuboff, *In the Age of the Smart Machine: The Future of Work and Power*, New York: Basic Books, 1988.

⁴Siegfried Giedion, *Mechanization Takes Command*, Oxford University Press, 1955.

⁵David F. Noble, *America By Design: Science, Technology and the Rise of Corporate Capitalism*, Oxford, UK: Oxford University Press, 1977.

force which lies at the heart of the growing concern over robotic weapons. What is clear is that automation and robotic technologies are poised to revolutionize the practices of modern warfare. The central question is what, if anything, the law can or should do to shape the coming revolution. This question leads us directly to the related questions of how new law ought to be formed, how an emerging technology might necessitate new law, and what the relation between the two ought to be. It is these questions which the current chapter seeks to address.

***Jus nascendi* in IHL**

Among the sources of new law, IHL has a diverse and inter-related means for establishing new law. From the recognition of customary law, to interpretive guidelines on existing law, to the doctrines and policies of states, to the precedents of courts and tribunals, to judicial opinions and decisions, to international declarations, treaties and conventions, and the interactions among all of these in establishing principles and norms, new international law comes into being in various ways. But it will be helpful to review some of the basic concepts and forms which international law takes.

Traditionally, international law is seen as coming from various sources and in varying degrees of abstraction—customs and norms, general principles and laws, written treaties and conventions, and jurisprudence. Customs and norms represent the traditional behaviors and expectations that can be observed in the decisions and actions of states, and from which general principles and laws may be derived. When it comes to written laws it is sometimes argued that conventions aim to codify customary law, putting into writing the norms of behavior already recognized and adopted by states. While this may be true of a broad range of international law, there are clear instances in IHL where written law has emerged specifically because the widespread behavior of states ran counter to shared moral sensibilities and collective interests. The Geneva Conventions of 1949 and the establishment of the United Nations stand as clear examples where the behaviors of states had failed to establish acceptable norms (as manifest in the horrors of World War II), and new prescriptive international law and regulating bodies were deemed necessary.

While customs and norms may be latent in the actions of states, they are often not described or formulated in a serious way until they become of matter of legal and political concern. Similarly, the derivation of underlying principles, including the writing of conventions, is usually motivated by some matter of diplomatic or legal concern. Among the proximate motivations to create new law we might identify various causes, including but not limited to: the emergence of new behaviors by states or non-state actors, the recognition or exploitation of ambiguities or inconsistencies in the existing law, or the emergence of new technologies that transform the capabilities, behaviors or effects of actions taken by states so dramatically that they undermine basic assumptions of customary or treaty law. In each case, the existing law can be seen as falling short in its purpose to protect the collective interests of states.

Underlying the whole of body of law, and particular expressions of law, we find morality. There are, of course, differing views on the relation of law and morality, as well as various legal and

moral philosophies. But there is a widely held belief that the law at least aims to express shared norms, and that many legal norms coincide with moral norms. Stronger views hold that the motivation and legitimacy of laws, and shared norms, derives from their basis in morality. It would be well beyond the scope of this paper to address the various perspectives on this foundational question in the philosophy of law. Fortunately, in the IHL context there is a longstanding legal clause that explicitly refers to morality—the Martens Clause. It is to a consideration of the nature of the Martens Clause, and its implications for creating new law that we now turn.

The Martens Clause

The Martens Clause is often invoked in disarmament contexts because it refers explicitly to the public conscience. It thus offers a more explicit role for public opinion and the representatives of civil society in the moral assessment of IHL. Many legal scholars, however, wish to resist an overly broad reading of the clause, and seek to limit its scope to a merely perfunctory recognition of customary law. Regardless of one's views on legal positivism, naturalism and integrity, it seems reasonable to view the law as a human construct which seeks to enact social values. The Martens Clause is not the only point where social and moral values weigh upon IHL—it is in many ways the whole point of IHL. Nor is the Martens Clause merely an acknowledgment that customs and practice can embody norms, even if they are not written down. A careful reading of the Clause and its history reveals something more interesting and complex than either of these accounts of it.

The Martens Clause first appeared in the Preamble to the Hague Convention II on The Laws and Customs of War on Land in 1899. It has been restated in various forms in numerous other conventions, most recently in the 1977 Additional Protocols to the Geneva Conventions. Originally proposed by Russian delegate to the Hague convention and jurist F. F. de Martens, its introduction was motivated by concerns over extending humanitarian law to armed partisans in occupied territories.⁶ According to Meron,

the clause has ancient antecedents rooted in natural law and chivalry. It is articulated in strong language, both rhetorically and ethically, which goes a long way towards explaining its resonance and influence on the formation and interpretation of the law of war and international humanitarian law.⁷

In its various formulations it became more explicitly clear that it expressed the notion that ethical consideration for the principles of humanity and the public conscience should provide a foundation for both customary and conventional law.

In its most recent formulation, the Martens Clause states:

⁶Theodor Meron, “The Martens Clause, principles of humanity, and dictates of public conscience.” *American Journal of International Law* (2000): 78-89.

⁷Meron, *op cit*.

Recalling that, in cases not covered by the law in force, the human person remains under the protection of the principles of humanity and the dictates of the public conscience.⁸ The most commonly held interpretation of the clause is that acts are not legal, or permissible, simply because they are not explicitly prohibited by the law. In recognition of the fact that many aspects of customary law are not explicitly written down, one cannot assume that something is legal simply because there is nothing in a given treaty or convention that prohibits it. This is a rather minimalist interpretation of the clause, however.

At the same time that it recognizes the status of customary law, the Martens Clause also acknowledges that the “principles of humanity and the dictates of the public conscience” are a valid source for new IHL. This is an acknowledgment of two parallel moral foundations for the protection of the human person within norms and customs. There are various ways we might unpack the concepts of the “principles of humanity,” and the “dictates of the public conscience,” and it is worth examining some of those ways.

The “principles of humanity” might be articulated as the principles derived from the UN Declaration of Human Rights. Alternatively, it might be interpreted to mean a more abstract set of moral principles which underwrite our conception of humanity, and were partially (and perhaps imperfectly) codified the UN Declaration of Human Rights itself. That is to say that there is no specific, written principle of humanity or set of principles, but these are evident in shared norms, and the Declaration of Human Rights explicitly sought to articulate them in a codified document. As such, those principles exist antecedently and independently of the document, but are also reinforced through that codification. As with other “general principles” of international law, the principles of humanity are probably better thought of as a hybrid of shared beliefs, sentiments, and attempts to derive principles from norms of belief and behavior, or codify them as declarations and conventions.

The “dictates of public conscience” might also be articulated in various ways. It is most often cast as being roughly equivalent to “public opinion,” but this is dissatisfying for several reasons. First, we must confront what we know about public opinion from the social sciences, as well as its being subject to manipulation by strategic communications and propaganda. But we must also recognize the morally relevant differences between public opinion and public conscience.

As a matter of expertise, the public is not always well informed of complex issues, such as international law, and likely does not hold any antecedent opinion on arcane matters with complex legal implications. This is not to say that only international lawyers should participate in determining the public conscience on such matter, but that they are likely the only ones with legal opinions on the subject. Indeed, since legal opinion can already be assumed to be relevant, what motivated the authors of the Martens Clause to include this phrase was to emphasize the moral, rather than legal, basis for norms. The point is that the dictates of public conscience cannot be reduced to legal opinions, and that moral conscience should not require special expertise.

⁸Additional Protocol II to the Geneva Conventions,

Treating the public conscience as public opinion ignores what we know about public opinion from the social sciences and communications research. Some have argued that public opinion holds unfair prejudices against autonomous weapons due to the negative portrayal of killer robots in science fiction books and film.⁹ While researchers have undertaken careful and objective studies of public opinion on killer robots and autonomous weapons,¹⁰ it is not clear that this is really the same thing as the public conscience that the authors of the Martens Clause had in mind. While at best public opinion might confirm what a more careful reflection on matters of conscience reveals, at its worse public opinion can be subject to manipulation and propaganda.

In presenting the public with pointed questions and simplified examples in the attempt to elicit their opinion, as is done in survey research, one can powerfully shape the expression of opinion. It is well known that the choice of words, and even the order of presenting questions can have significant impacts on the responses. Critical studies of public opinion have demonstrated the concept itself to be a social construct that serves primarily to coordinate political communications.¹¹ The communications industry has deployed the social sciences so as to deliberately construct publics whose opinions can be easily categorized and influenced. This has become an art and science for public relations professionals, and its influential power can be seen clearly in commercial advertising, as well as in contemporary U.S. elections. In its most extreme forms we call this manipulation propaganda,¹² but even in its milder forms it belies the fact that public opinion is not only dynamic but malleable and controllable to some extent. As such, we ought to question its capacity to serve as a foundation for new law.

More fundamentally, “conscience” has an explicitly moral inflection which “opinion” lacks. This implies not merely holding an opinion or belief on a moral issue, but actually feeling compelled by, or believing in, a specific moral obligation or duty. That is to say, one may feel the weight of moral conscience even when one acts or believes against it (we may even feel it most acutely when violating it, as regret). Thus, moral conscience is not simply a result of a moral deliberation—the choice of action, or ultimate belief or opinion about which action is appropriate.

⁹Charli Carpenter, “The SkyNet factor: Four myths about science fiction and the killer robot debate,” *The Washington Post*, September 3, 2014. Downloaded from: <http://www.washingtonpost.com/blogs/monkey-cage/wp/2014/09/03/the-skynet-factor-four-myths-about-science-fiction-and-the-killer-robot-debate/>

¹⁰Charli Carpenter, “How scared are people of “killer robots” and why does it matter?” *Open Democracy*, July 4, 2014. Downloaded from: <https://www.opendemocracy.net/charli-carpenter/how-scared-are-people-of-%E2%80%9Ckiller-robots%E2%80%9D-and-why-does-it-matter>

¹¹Nikolas Rose and Thomas Osborne, “Do the Social Sciences Create Phenomena: The Case of Public Opinion Research,” *British Journal of Sociology*, 50, 3 (1999): 367-396.

¹²Edward Bernays, *Propaganda*. Routledge, 1928.

Moral conscience includes the felt forces of duty and obligation, and the moral sentiments attached to the processes of moral deliberation in their totality.

As such it is a disservice to reduce the “dictates of public conscience” to mere public opinion, even if we might use the tools of the social sciences to develop insights into the contents of the public conscience. That content should also be elicited through public discussion, as well as academic scholarship, artistic and cultural expressions, individual reflection, collective action, and additional means, by which society deliberates its collective moral conscience. Indeed, the best place to look for emerging norms and the dictates of public conscience are in the public forums in which states and individuals attempt to grapple with, and articulate that conscience. Which is one reason why it is so crucially important that the representatives of civil society have a voice in those forums, and that those voices represent a broad array of national and individual perspectives.

A Nascent Principle - Meaningful Human Control

I want now to examine the recent debates and discourse surrounding autonomous weapons with an eye toward the nascent norms and principles we might find there. Three key sets of concerns have emerged from the discussion of whether to regulate autonomous weapons systems. The first set are the concerns over the risks to civilians that such weapons pose. The second set of concerns are the implications for human rights and human dignity from the use of such systems. The third set are the concerns over how autonomous weapons might transform the use of violent force in warfare in terms of the psychological responsibility taken by humans in decisions to use force and the legal accountability for intended and unintended consequences for the use of such systems. Each set of concerns is warranted, as autonomous weapons systems potentially pose serious risks to civilians, to human rights and dignity, and to the laws and norms of accountability. And while each set of concerns implies a need for scrutinizing the development and use of autonomous weapons systems, they do so somewhat differently and suggest different types of norms.

On its surface, the concerns over the risks to civilians makes autonomous weapons out to be analogous to landmines or cluster munitions in terms of the potential of these weapons to have significant and unintended impacts on civilians and civilian property. This concern leads many to cast the central question as whether or not autonomous weapons can be used in conformity with existing IHL requirements. In particular there is a concern as to whether autonomous weapons can conform to Article 36 of the 1977 Additional Protocol I of the Geneva Conventions of 1949, and its requirements that new weapons systems must be capable of being used in a discriminate and proportionate manner. That is, if a weapon is considered to be *intrinsically* incapable of discriminate and proportionate use, it ought not to be fielded or used. Any weapon could potentially be used in an indiscriminate or disproportionate manner, so that potential is not sufficient to prohibit its use. For those weapons in which there are circumstances in which their use would be indiscriminate or disproportionate, their use ought to be restricted and their use

should only be permitted in those contexts and circumstances where they can be used in compliance with IHL.

Unguided missiles are a common example of weapons which can be assessed as discriminate or indiscriminate depending on the circumstances. Lacking guidance systems, it is considered that such missiles cannot be used discriminately in populated areas.¹³ They are not prohibited altogether, however, as there are contexts in which they might be used where they do not pose any risk to civilians, because there are no civilians in the vicinity of the targeted area. Of course, in a highly restricted environment, meeting sufficient conditions, (e.g., a large enough area occupied exclusively by enemy combatants and their equipment) then it may be possible to claim that any weapon could be used with some degree of discrimination and proportionality. This was, in fact, the argument made in defense of the acceptability of cluster munitions—they could be used in certain ways in certain situations in which they do not have indiscriminate or disproportionate effects. The argument ultimately turned not on the context of initial use, but on evidence from the actual use of these weapons, which demonstrated the persistent effects of unexploded munitions on civilians, and the ongoing dangers that they posed long after an attack and even long after the conflict.

Many of the critics of disarmament proposals directed at autonomous weapons argue that the risks to civilians are unknown, and that there may be benefits to civilians in terms of lowered risk if such technologies are allowed to be developed.¹⁴ These critics are right to identify the

¹³See: International Committee of the Red Cross, *Customary IHL*, Rule 12 “Definition of Indiscriminate Attack” and Rule 71 “Weapons that are By Nature Indiscriminate”:

https://www.icrc.org/customary-ihl/eng/docs/v2_rul_rule12,

https://www.icrc.org/customary-ihl/eng/docs/v2_rul_rule71

Maya Brehm (2014) “Unacceptable Risk: Use of explosive weapons in populated areas through the lens of three cases before the ICTY,” PAX Report, October, 2014 pp. Downloaded from: <http://www.paxvoorvrede.nl/media/files/pax-rapport-unacceptable-risk.pdf>

¹⁴See: Ronald C. Arkin, *Governing Lethal Behavior in Autonomous Robots*, CRC Press, 2009; William Boothby, *Weapons and the Law of Armed Conflict*, Oxford University Press, 2009; Gary Marchant, Braden Allenby, Ronald C. Arkin, Edward T. Barrett, Jason Borenstein, Lyn M. Gaudet, Orde F. Kittrie, Patrick Lin, George R. Lucas, Richard M. O’Meara and Jared Silberman, “International governance of autonomous military robots,” *Columbia Science and Technology Law Review*, 30 December 2010, available at: <http://ssrn.com/abstract=1778424>; Kenneth Anderson and Matthew C. Waxman, “Law and ethics for robot soldiers,” *Policy Review*, 28 April 2012, available at: <http://ssrn.com/abstract=2046375>; Michael N. Schmitt and Jeffrey S. Thurnher, “Out of the Loop: Autonomous Weapon Systems and the Law of Armed Conflict,” *Harvard National Security Journal*, 231 (2013); Marco Sassoli, “Autonomous Weapons and International Humanitarian Law: Advantages, Open Technical Questions and Legal Issues to be Clarified,” *International Law Studies / Naval War College*. 2014, vol. 90, p. 308-340; Kenneth Anderson, Daniel Reisner and Matthew Waxman, “Adapting the Law of Armed Conflict to

importance of the norms and conventions that require the discriminate and proportionate use of force, and the protection of civilians. They are wrong, however, insofar as they attempt to limit the concerns over autonomous weapons to only these norms.

There is an extreme view that the only purpose of IHL is to try to reduce the risk to civilians in war. But the various provisions of IHL that apply only to combatants—from rules regulating surrender and prisoners of war to the treatment of deceased combatants—make such a view untenable. A slightly less extreme view holds that the goal of IHL is to protect an essential core of humanity, even amongst the great inhumanity of war—which brings us back to the “principles of humanity” in the Martens Clause. A broader view holds that IHL has many aims, not all of which are reducible to risk minimization or preserving the principles of humanity, but that those aims must not supercede the principles of humanity.

Moreover, as I have argued before,¹⁵ existing IHL imposes specific requirements on decision makers, who are at least implicitly human. It is at best ambiguous what it would mean to replace that decision making with non-human computers and program—even though they are creations of humans. We may, in fact, be morally and legally prohibited from delegating such decision-making to automatic systems. At the very least, existing customs and norms, as well as written conventions, governing decisions to use violent force all assume human decision making. These norms, and the interpretation of existing laws, are called into question when human decision making is replaced with automated processes. The norms that govern the requirements of taking responsibility for command decisions, and holding individuals and states responsible for the consequences of their decisions, are certainly a matter of concern in the development and use of autonomous weapons. In order for the decisions to use force to count as legal decisions, or moral decisions, they must be the considered judgements of a human in a given situation, assessing the available information. An automated process designed in advance, based on sensor data, is neither a legal judgement nor a moral judgement. Similarly, rules of engagement are not decisions to engage or use force—rather they are guidance to human decision makers who will ultimately make those decisions and carry the responsibility for them.

When thinking about new law, it is helpful to consider the nature of norms and how they inform law. This is somewhat challenging when we consider new technologies, especially those which have the potential to greatly disrupt existing practices. Philosophers view norms as concepts which imply obligations on actions—permissions or prohibitions on certain actions. In the social sciences, social norms are conceived as patterns of social behavior which, in virtue of being reliably observable patterns, manifest an underlying set of shared beliefs about the acceptability of certain forms of behavior. International law combines the concept of a belief to act according to obligation together with these observable patterns of behavior. Thus the standard view of customary law is that customs must be recognized in the consistent actions of a significant

Autonomous Weapons,” *International Law Studies*, forthcoming 2014.

¹⁵Asaro, *op cit*.

number of states based on beliefs that there is an obligation to so act, and not be rejected or ignored by a significant number of states.

More specifically, Article 38 of the Rome Statutes of the International Court of Justice establishes four sources of international law: conventions, customs, general principles, and judicial decisions. Article 53 of the Vienna Conventions on the Law of Treaties further distinguishes a special class of norms as peremptory norms—those which cannot be derogated.¹⁶ Of course, there is no definitive list of which norms are peremptory, nor is there a definitive list of general principles, or of customs. Conventions are generally conceived of as written codifications of customs and general principles. As such, they do attempt to define norms and principles, though they recognize their own limitations in doing so, as the Martens Clause itself makes clear. Treaties are seen as instruments which impose restrictions on the behavior of states which would not otherwise be restricted under custom or convention, but are limited in that they still cannot derogate peremptory norms.¹⁷

In its most abstract form, technology is simply a system of skills, tools and practices.¹⁸ The introduction of any new techniques, skills or technological tools necessarily implies a transformation in practices, whether at the level of individual tasks or at the levels of task management or institutional organization. The question is whether these transformations conform to existing norms, or challenge those norms. More precisely, there are four possibilities: 1) Existing norms may continue to govern practices after the introduction of a new technology, largely unchanged.

- 2) Existing norms may be diminished or dissolved after the introduction of a new technology.
- 3) New norms may emerge to govern the practices after the introduction of a new technology.
- 4) Existing norms may be transformed or reconceived.

When considering the norms governing new technologies, we are presented with a chicken-and-egg problem. If we have not yet implemented a new technology, we cannot observe what the new

¹⁶Among the generally accepted peremptory norms are prohibitions on genocide, torture, slavery and piracy.

¹⁷Thus, a group of States could not form a treaty organization to permit themselves to conduct genocide, torture, slavery or piracy—or at least such a treaty would not be accepted as removing those states from respecting the international norms prohibiting these.

¹⁸In ancient Greek *techne* meant the performance of a craft, what in modern English we might call technique. Of course the performance also requires the use of the body and tools, and in modern English we often use “technology” to refer to sophisticated tools. Yet with sophisticated technologies, such as computers, we are really concerned with their performance, which entails the practice of skills as well as the functionality of the tools. Thus, we should think of technologies as consisting of the function of material tools and the skilled practices of people in conjunction, as well as their systematic organization.

norms are. We can examine existing norms, and try to determine if the use of a new technology would challenge or violate those norms. If so, we might try to regulate that technology, and try to ensure that the norm remains in effect. Or we might find that it does not violate those norms, and so we should not need to regulate it, and the existing norms ought to be sufficient until proven otherwise. But there are further possibilities. We might find, after a technology is adopted, that new sets of practices emerge, followed by new norms.¹⁹ We may also find that the capabilities of a new technology actually manifest situations in which we recognize norms that had always been tacitly assumed, but never articulated or codified because it was previously unnecessary.

Rather than framing the current debate over autonomous weapons as a utility calculation which weighs the potential risks of a new technology against its potential benefits, it would be better to view it as a question of how we might best regulate this new class of technologies, which might take many forms and have various sorts of capabilities which challenge existing norms. That is, while autonomous weapons do pose many obvious risks, it would be difficult to regulate or mitigate these based on capabilities which are as yet unknown.²⁰ Rather, it makes sense to focus on the norms that will be challenged or violated by this new technology, and try to determine how best to protect those norms from being undermined. Because of the new capabilities these technologies will bring, it is also important to consider what nascent norms may be in place which have not previously been articulated because the need had not previously been questioned.

In the case of autonomous weapons, what is really new is the automatic selection of targets and decisions to use violent force against those targets. While Sharre²¹ and others have pointed out that there are various weapons systems already in use that have simple implementations of these capabilities, it is not at all obvious that the practices and norms governing the use of current systems will adequately scale to the sophistication and complexity that robotic and autonomous weapons systems appear poised to achieve in the coming years and decades. Even Anderson & Waxman²² agree that new practices and norms must emerge to govern them, but would rather leave this process to self-regulation and the dissemination of best practices and soft law. So the debate is not that new norms will emerge, but where they should come from, how they should be articulated, and whether or where they should be codified.

In discussing these issues in various forums over the past few years, what seems increasingly

¹⁹Among these are “best practices” and other forms of soft law.

²⁰See Note 10 above.

²¹Paul Scharre, *Robotics on the Battlefield, Part I: Range, Persistence and Daring*, Center for New American Security Report, pp. 47. Downloaded from: http://www.cnas.org/sites/default/files/publications-pdf/CNAS_RoboticsOnTheBattlefield_Scharre.pdf

²²Anderson and Waxman (2012), *op.cit.*

clear is that there are some nascent norms and principles regarding autonomous weapons which are widely shared, and which are beginning to take shape. What emerged from the discussions at the Experts Meeting convened by the Convention on Certain Conventional Weapons at the United Nations in Geneva in May of 2014,²³ was a growing convergence towards a new norm or principle, that of “meaningful human control.” There remains some disagreement and confusion, at least in terminology, over the various definitions of automatic, autonomous, and the nature of intention, agency and responsibility in highly automated and autonomous systems. Yet despite this terminological disagreement, there was broad based agreement that military attacks and the use of violent force in war should always be kept under “meaningful human control.” What exactly constitutes meaningful human control, and how we might evaluate whether a given weapons system conforms to a possible requirement for it, are matters that remain to be articulated.

However, it seems quite clear that we have something that looks very much like an emergent principle. It is difficult to call this a norm proper, even though it is the case that states currently do act so as to keep weapons under meaningful human control. It is not as clear that they actually believe that they have a shared obligation to do so. That belief and obligation might emerge in the future as practices develop around new autonomous weapons technologies. But such a norm might not emerge in practice in the absence of regulations or guidelines. The real risk here, from a legal perspective, is that practices based in political and military expediency might give rise to weak norms, or no norms at all, and the use of weapons without any meaningful human control may become acceptable as a *de facto* norm. In that case, we might find ourselves in a world where the risk of unintended consequences (or the ability to strategically obscure one’s intentions) in using such weapons is substantial, while the responsibility for such use by states and individuals is limited or easily avoidable in practice.

Some legal scholars have argued that no professional militaries would want to use such weapons, because they would have unpredictable results. But this is rather speculative, as such systems are not yet available, and the behavior and beliefs of states and their militaries may change as new technologies begin to enable new kinds of operations, tactics and strategies. Once states begin to see some advantage to using such weapons, they may be reluctant to impose any regulations, and in the event of widespread use there may be little basis in customary law to reign them in, even if a handful of countries with advanced professional militaries do observe a stricter set of norms, while a significant number do not.

As a matter of analogy, I believe that it is helpful to look at the concept of “superfluous injury and unnecessary suffering” in IHL. This phrase first appeared in the St. Petersburg Convention of 1868. At the time, it was considered a norm, but one which had not been articulated clearly before. That is, militaries already recognized that one should not intentionally cause injury or

²³See: CCW website on Lethal Autonomous Weapons, [http://www.unog.ch/80256EE600585943/\(httpPages\)/6CE049BE22EC75A2C1257C8D00513E26?OpenDocument](http://www.unog.ch/80256EE600585943/(httpPages)/6CE049BE22EC75A2C1257C8D00513E26?OpenDocument)

suffering beyond what was necessary to achieve a military objective. So while it is permissible to kill or wound in order to achieve an objective, adding additional injury or increased suffering which serves no military purpose ought to be avoided. It could be argued that the norm had existed all along, but it was the introduction of new technologies—exploding and fragmentary bullets, in that case—which inspired the articulation and codification of this normative principle.

In some sense the prohibition on the use of these new types of bullets, and the recognition of this nascent norm emerged together. In another sense, the recognition of the norm and its underlying principle were fomented by unease at what these new weapons represented for the future of warfare. That is, we might consider that there was something in the moral sensibility and conscience of the delegates to the St. Petersburg convention that the use of such weapons is morally wrong, but also that they may not have been able to state clearly exactly what was wrong about it before attempting to find the appropriate language for it. It took some effort to work through what their moral conscience was, and to express this in words, as well as to reach consensus on how to codify this in law, and to articulate the prohibitions on weapons that it implied.

Thus, the principle of meaningful human control would appear to be something that has historically been taken for granted—assumed but never stated. Weapons always required humans to decide when and where to use them. While booby traps and mines challenged this implicit assumption to some extent, it was not sufficient to motivate the articulation of a new principle, though it did spawn various practices, norms and even treaties. Autonomous weapons pose a much greater challenge to our previous assumptions about the determination of targets and decisions to use violent force. As such, they have motivated a discussion and reflection on the principles of humanity and the dictates of public conscience. It is from this reflection that the principle of meaningful human control has emerged. And indeed, given the uncertainty over how such systems might be developed in the future, it is all the more important to clarify the fundamental principles.

Meaningful human control, as it has thus far been articulated, contains several elements.²⁴ First, it

²⁴The phrase “meaningful human control” was coined by Richard Moyes, and its initial articulation first appeared in the briefing papers of the UK-based NGO Article 36: Article 36 (2013) “Killer Robots: UK Government Policy on Fully Autonomous Weapons,” Briefing Paper, April, 2013, 5 pp. Downloaded from: http://www.article36.org/wp-content/uploads/2013/04/Policy_Paper1.pdf
Article 36 (2013) “Structuring Debate on Autonomous Weapons Systems,” Briefing Paper, November, 2013, 3 pp. Downloaded from: <http://www.article36.org/wp-content/uploads/2013/11/Autonomous-weapons-memo-for-CCW.pdf>
Article 36 (2014) “Key Areas for Debate on Autonomous Weapons Systems,” Briefing Paper, May, 2014, 4 pp. Downloaded from: <http://www.article36.org/wp-content/uploads/2014/05/A36-CCW-May-2014.pdf>

is fundamentally humanist in its insistence on explicitly human control of targeting and firing decisions. If any new principle might be convincingly derived from the “principles of humanity” as expressed in the Martens Clause, surely it would be a principle that ensures human control over the violence of war, and war itself.²⁵

It is not yet clear what constitutes control exactly. So it seems that more discussion of this would be prudent. While a strict definition is not necessary, a shared understanding should be a goal of further discussion.²⁶ At the very least, control implies that the effects and potential consequences of using a weapons system must be reliable and predictable to an extent that the human can exert some form of control over it. It also implies that the performance of the system must conform to the intentions of the operator such that it is possible to distinguish when a system is under control, or when an operator has lost control.

Another key aspect of meaningful human control is the “meaningful” part. While it is difficult to define exactly what meaningful control consists of, the concept itself is clear and it seems a reasonable standard, if not an obvious one. In part it aims to prevent weapons systems that use humans instrumentally as approval mechanisms. For example, ordering a soldier to press a “fire” button every time a light comes on would imply that a human is technically “in control” of the weapon system even if they effectively have no meaningful control over what the system is targeting, or how and when it is using lethal force against those targets. But it also implies more than this, in that meaningful control also entails taking responsibility for the use of the weapon system, and being accountable for the consequences of that use.

In order to ensure that a system retains human control, and thereby upholds the principles of humanity and does not trivialize human actions in the process, the requirement for control to be meaningful is necessary. For the killing of a human to be meaningful, it must be intentional. That is, it must be done for reason and purpose. Philosophically, intentionality requires understanding the meaning and significance of an act. While autonomous systems may be programmed to act in a certain way, given a certain set of conditions, they cannot understand the significance of their acts. This is in part why they cannot make legal or moral judgements. But this also relates to the question of human dignity. If a combatant is to die with dignity, there must be some sense in which that death is meaningful. In the absence of an intentional and

²⁵Mark Gubrud, “The Principle of Humanity in Conflict,” ICRAC Blog, November 19, 2012. downloaded from: <http://icrac.net/2012/11/the-principle-of-humanity-in-conflict/>

²⁶Consider, again by analogy, the principle that prohibits “superfluous injury and unnecessary suffering.” These seem fairly clear as a matter of principle—injuries and suffering that go beyond military necessity. Yet as a matter of practice, or judging individual weapons, it is far from clear how one might apply this principle to a new weapon under review. How does it apply to pain-inducing weapons that do little physical damage, for instance? In practice, of course, states have developed norms in the application of the principle, shoring it up with more explicit treaties or guidelines as necessary.

meaningful decision to use violence, the resulting deaths are arbitrary and their significance , along with the dignity of those killed, is diminished.

Meaningful human control also offers some positive guidance on how systems ought to be designed to interface with humans. It obliges engineers and designers to consider how the use of violent force is a form of symbolic and intentional action, as well as a functional performance. As such, an interface ought to provide its user with the potential to make meaning and take meaningful actions, as well as to perform tasks. Moreover, it couples the making of meaning with control over the system—to the extent the system is automatic, the meaning of its activities are dependent upon higher order levels of organization, which may not be able to address what is significant in a given situation. Meaning-making is a distinctly human capacity, and artificial systems will lack such capabilities for the foreseeable future. The use of anthropomorphizing language can often confuse the real capabilities of systems, which is why it is imperative that we make principles like meaningful human control clear.

Conclusion

I hope that I have shown not only that we should view international law as an evolving and dynamic system, but also that its evolution ought to be shaped by moral considerations. The Martens Clause ought to be viewed not only as a recognition that written law does not supplant customary law, but also as an invitation to moral reflection on the role of the principles of humanity and the dictates of public conscience in articulating and establishing new IHL. Its legacy as an explicit recognition of the role of moral consideration in the application of IHL, as well as in the formulation of new law should not be overlooked or underestimated.

And finally, I presented as an example of an emerging normative principle concerning the development and use of autonomous weapons—the principle of meaningful human control. While it is in some sense nascent or latent in existing beliefs and practices, it is also in some sense emergent in the ongoing debates that are themselves a response to the expectations and capabilities of emerging technologies.

It is my hope that the current debate over the regulation of autonomous weapons can progress beyond speculative assessments of the capabilities and risks of robotic technologies. Instead, we should focus on the threats posed to fundamental norms of responsibility and accountability, and to the threats to human rights and human dignity that these new technologies represent. And we should draw upon the principles of humanity and dictates of public conscience as we move forward in articulating new normative principles and *jus nascendi*, including the meaningful human control over the use of violent force in armed conflict.