The Role of the Ethical Underpinnings of International Humanitarian Law in the Age of Lethal Autonomous Weapons Systems

Abstract: This paper presents selected conclusions related to the theoretical underpinnings of international humanitarian law, with special focus on the understanding of considerations of humanity and the dictates of public conscience (the Martens clause) and their impact on the regulation of lethal autonomous weapons systems. Despite the fact that different positions can be found in the doctrine, it is argued herein that the general principles of international humanitarian law are not sufficient to properly regulate the disruptive military technologies (new means and methods of warfare) and a new international norm is needed. Consequently, the paper agglomerates extra-legal and cross-cutting arguments stemming from other normative regimes that point to prioritization of the value of human life and the role and quality of the human factor in decision-making procedures relating to the health and life of victims of modern armed conflicts, which should be incorporated in it.

Keywords: armed conflict, artificial intelligence, humanity, Martens clause, autonomous weapons

Introduction

The current discussion about the legality of new military technologies is marked by two issues with are both somehow opposed and connected. One can easily find an argument raised by States and subject matter experts that the current paradigm of international humanitarian law (IHL), that is the balancing of considerations of humanity and military necessity, is sufficient and always operative (Boothby, 2016; Schmitt, Thurner, 2013; Wagner, 2014). But at the same time, since this argument is not fully convincing there is a growing...
The Role of the Ethical Underpinnings

recognition of a need to refresh and bolster the theory underpinning this predominantly practical branch of law (Arkin, 2009; Asaro, 2012; Crootof, 2016; Liu, 2012).

The aim of the reflections contained in this paper is to present selected arguments that have become consolidated within the discussion on lethal autonomous weapons systems (LAWS), and more precisely to present selected theoretical and ethical underpinnings of IHL and confront them with the development of artificial intelligence equipped with deadly force.

For the purpose of the following considerations, LAWS shall be operationally defined as military combat systems that can select and engage human targets without meaningful on-hands human control, i.e. without conscious human-led decision-making. In this context I test the claim that there is no reason to panic (Jenks, 2016) because there are general principles and concepts of IHL which sufficiently regulate the use of LAWS. This claim is built upon the fact that IHL has been dealing for centuries with new technologies and another tech-revolution does not require a normative evolution (Haines, 2014).

Taking the above into consideration, I try to compactly present my arguments in the debate on whether the robotization of warfare is leading to a shift in the treatment of human beings in an existential context. I do not agree with the adequacy of general principles of IHL in the era of LAWS, as their introduction on modern battlefields is qualitatively different than previous revolutions in military technology (such as anti-personnel mines or blinding lasers), precisely because of the change in the approach to the value of human life. Thus this work concentrates on the content and function of the paradigm of IHL and the role of extra-legal considerations introduced via the Martens clause.

The Eternal Unknown in Balancing Humanity and Military Necessity

The starting assumption of my research is that IHL constitutes an undertheorized legal regime and lacks theoretical models capable of providing a systemic interpretation. IHL is a highly pragmatic set of rules which, according to Kolb, form „in itself a law of emergency or necessity” (2014, p. 85) that in its imperfection benefits from a high level of confidence shared by both military commanders and humanitarian activists. This is due to its balanced nature. It incorporates humanitarian considerations, embodied in the so-called principle of humanity, as well as acknowledges that military necessity provides a license to use military force in justified instances.

The Momentum for the Balancing Exercise

When it comes to the conduct of hostilities, this balancing exercise is presented as a spirit or fundamental paradigm (Hayashi, 2016a; Kleffner, 2012). However, such a perception is insufficiently instructive as it usually does not refer in any way to the concrete impact of balancing humanity and military necessity on the creation, interpretation, or application of
IHL. To the contrary, it may even be perceived as a source of confusion, which starts with highly theoretical doctrinal discussions (Kalshoven, 2007; Kleffner, 2012; Goodman, 2003; Parks, 2010; Schmitt, 2013) and may end with a single commander and his legal adviser being trapped in the obscurity of current legal developments or ever more frequently the commander left alone to decide.

An example that best illustrates this problem is the concept of „least harmful means”, which according to Grzebyk (2019) divided law experts into two factions: the „capture faction” (e.g. R. Goodman, J. Pictet, A. Szpak) and the „kill faction” (e.g. L.R. Blank, Ch. Jenks, E. Talbot Jensen, M. Schmitt, J.D. Ohlin, W.H. Parks or J. Kleffner). Upon a closer examination of the arguments of both factions, one can put forward the thesis that a part of the doctrine supporting the introduction of a gradual use of force (the „capture faction”) is based on an interpretation of considerations of humanity and military necessity that goes beyond the level that has already been incorporated into the norms of IHL. Thus the ‘necessity’ to save the lives of combatants is derived from an overinterpretation of the principle of distinction and proportionality in the context of considerations of humanity (resulting from the concurrent application of human rights law). Consequently, the view of the traditional faction (the „kill faction”) takes into account the balancing, previously incorporated by lawmakers, of both considerations.

As the doctrinal discussions continue, one thing is certain – the impact of the ambiguity of the balancing exercise on the application of IHL does not contribute to the improvement of security of the actors involved. And since States rarely take an official position (expressed in a norm-creation exercise for the balancing between humanity and military necessity), issues such as the increased standard of protection of one's own soldiers, the need to arrest enemy combatants instead of killing them, the obligation to offer surrender terms, the obligation to use the most discriminatory and most precise weapons (including replacing soldiers with weapons without operators) will continue to raise doubts and will be discussed in the doctrine without any real transposition into the battlefield.

As a result of the above ambiguities it can be argued that this crucial – and at the same time insufficiently analyzed – balancing exercise can be undertaken on two different levels. Some commentators contend that it should be performed it in every application of IHL, while others contest this extensive approach and conclude that a strict interpretation is imperative, as the configuration of the balance between the two considerations depends on the lawmakers and the particular situations in which they wish to permit/tolerate a specific course of action.

I acquiesce to the latter view and would add that the ongoing robotization of warfare presents a new opportunity to reshape that balance according to the modern understanding of the values of human dignity and humanity in the context of LAWS and the resulting qualitative shift in warfare. The opposite view would require a double configuration of the balancing, first performed at law-making level, and secondly on the occasion of implementing and enforcing the law. Such an approach could lead to new and even contra legem find-
ings, which could contravene the will of lawmakers. Therefore, it should be agreed that this crucial balance between humanity and military necessity is already incorporated in the legal norms of IHL and should not be changed or redefined depending on other factors, like the political, cultural, or ethical predispositions of the States engaged in an armed conflict.

The Content of the Balancing Exercise

The foregoing discussion entails a logical question which is critical to the whole debate: What is in fact the subject of this balancing – principles, rules or extra-legal values? In other words, what is the gist of humanity and military necessity?

Although the „principles” of humanity and military necessity do not appear as such in the IHL treaties, they constitute an inseparable element of the exegesis of the IHL norms, constituting both their justification and basic considerations. Considerations of humanity and military necessity also stem from such original norms as avoiding unnecessary suffering, useless destruction, and the obligation of humane treatment (Żeligowski, 2014). Humanitarianism is a concept which in most cases is applied in an intuitive way, without giving a strict definition, via reference to values which, at least at the general level, are considered as universally shared (Pictet, 1956). While invoking the considerations of humanity, lawmakers and the actors involved in the international discourse must relate to a universal meaning reflecting moral or even existential values, which can be expressed, inter alia, through the dictates of public conscience. Nevertheless, despite its semantic richness (or perhaps because of it), it is difficult to indicate a specific definition of the „principle” of humanity. It is more convincing to perceive it, in accordance with the ICJ jurisprudence (Legality of the Threat or Use of Nuclear Weapons, 1996) in the light of the other principles of IHL (i.e. the principles of distinction, proportionality, and the prohibition of unnecessary suffering and superfluous injury) which in a complementary way implement the normative ideal enshrined in the „principle” of humanity. This is indicated by the inclusion of considerations of humanity in the Geneva Conventions (GC) and Additional Protocol I (API) as the „laws” and „principles” of humanity in the plural form, and the „laws of humanity” in the preamble of the Fourth Hague Convention (Art. 63 I GC, Art. 62 II GC, Art. 142 III GC, Art. 158 IV GC) and „the principles of humanity” in Art. 1 point 2 API, as well as in the preamble to the CCW Convention, the Ottawa Convention, the CCM Convention, and the Treaty on the Prohibition of Nuclear Weapons.

Military necessity, like considerations of humanity, also occupies a fundamental place among the normative bases that guide the activities of States involved in armed conflict. Military necessity allows what is necessary to be done, while what is not necessary from a military point of view is prohibited. Nowadays military necessity is no longer considered a standard allowing all activities necessary from a military point of view (Kriegsräson), and the provisions of IHL set limits on the acceptable level of violence.

In the balancing exercise, military necessity is understood more broadly than just military necessity giving permission for a given action within the framework of a legal norm
(a type of a license). It provides for a purely utilitarian approach to armed actions, since its purpose is to recognize the belligerent’s interest in achieving victory while ensuring it uses only the military resources necessary, i.e. uses as few such resources as possible. Therefore, military necessity should correspond to a sense of equity (understood as a method to reduce the humanitarian losses and the use of military resources) from the point of view of the military decision-makers responsible for the correct conduct of combat operations.

Since considerations of humanity and military necessity provide the determinants in the process of lawmaking and application of the law, it is legitimate to adopt a broad interpretation of them. This translates into going beyond the classical and narrow understanding of them as an implementation of the prohibition of unnecessary suffering and superfluous injury (humanity) while providing a license provided for in IHL law norms (military necessity). In other words, the principles provide incentives to adopt new regulations and to reinforce argumentation when interpreting specific provisions (Hayashi, 2016b).

Hence, the International Committee of the Red Cross (ICRC) emphasizes that both considerations cannot repeal or replace the specific provisions of IHL, but constitute „guidelines” for interpreting the rights and obligations of the belligerents established by these provisions (ICRC, 2009). While it is not an uncommon assertion that the value of IHL principles is attributable to these basic principles, this argument however is not convincing (Dinstein, 2012; Schmitt, 2010). As rightly pointed out by Kleffner (2012), neither principle has a legal character nor the standing of principles or rules of law on their own. Instead they are factors to be taken into account by States in the lawmaking process, during which their balancing occurs. Thus, in order to distinguish them from the general principles and rules of IHL, they should be presented and understood as the primary considerations of this branch of international law (therefore not as legal norms per se).

In view of the above, this analysis assumes that both concepts have a dual function: first and foremost, they constitute lawmaking factors which belong to the ‘incentive zone’ of subjects of international law, and according to Art. 31 of the Vienna Convention on the law of treaties they should be taken into account in the process of interpretation and clarification of a given legal norm. This should mean that that the provisions of IHL shall be interpreted in the light of their object and purpose, which is the balancing of humanity and military necessity already enshrined in the given provision. This in turn means that to correctly implement the law and fulfil its purpose, one would need a specific treaty or customary norm or general principle of law on which to base his or her judgement and/or action, since in such a case a simple reference to the considerations of humanity and military necessity would not be sufficiently instructive.

This is precisely the case of LAWS and other disruptive technologies, the legality of which cannot be assessed in abstracto with sole reference to considerations of humanity and military necessity. They should rather be regulated in a specific norm of international law (treaty or custom). While the general principles of the IHL should be applied in a complementary way, it is argued here that they are not sufficient because they do not relate to the
The Role of the Ethical Underpinnings

qualitative shift in question, i.e. the transfer of the decision-making process about people’s lives to artificial intelligence.

Also, it should be noted that there is no example of a means or method of warfare that has been regulated by a customary norm or general principle of law without a separate regulation in treaty law.

As of today, a specific customary or treaty norm on the legality or use of LAWS does not exist, but the discussions taking place at the CCW Convention in Geneva since 2013 have at least provided a multi-disciplinary forum for „brainstorming“, which consequently makes it possible to disseminate some interesting ideas on the norms that may be applied, or at least may serve as a source of inspiration for States to eventually accept a new legal norm relating to autonomy in warfare.

Ethical Conduct of Armed Conflicts

The basic principles of IHL expressing the considerations of humanity share the pursuit of congruous goals: the protection of human life and respect for humanity in situations of armed conflicts, while they acknowledge the necessity of killing people. Humanitarianism in combat operations refers to the efforts of the belligerents to mitigate the suffering that is caused during armed conflict by providing victims, insofar as possible, with protection and assistance. These standards are of course well-framed in legal norms of IHL, and in addition they are also strengthened by the extra-legal standards set forth in the Martens clause and can be supported by legal arguments stemming from the regime of administrative law and human rights law, which provide norms for the proper execution by a State of its power and sovereignty.

The Martens Clause as a „Safety Valve”

The Martens clause is often presented as the final „safety valve” to be activated in the absence of clear legal norms in formal terms, because it constitutes the essence of the IHL regime. The case of LAWS seems to fit this hypothesis perfectly.

Kolb (2014) recognizes that the Martens clause (together with military necessity) provides for basic mechanism for classical IHL to be applied by States in areas that have not been regulated by law. While military necessity as the key component of the law of war has lost its importance with the adoption of the Geneva Conventions, the Martens clause has retained its status as a standard which prohibits the mechanical application of the concept that „what is not expressly forbidden is permitted” (Kolb, 2014, p. 8). The Martens Clause is the cornerstone of the IHL, understood as a set of norms restricting the freedom of sovereign States to act, both in times of war and peace (The Fourth Hague Convention preamble, Art. 63 I GC, Art. 62 II GC, Art. 142 III GC, Art. 158 IV GC, Art. 1 point 2 API, The CCW Convention preamble). Despite its controversial nature (Cassese, 2002; Dinstein, 2012; Meron, 2000) it
can be argued that it has a practical input in two possible areas: as a reference to extra-legal considerations in the law-making process, and in the decision-making process.

**A Fresh Perspective on the Dictates of Public Conscience**

The state of discussions at the CCW Convention and the number of speeches in which experts dealing with legal, ethical, and technological issues refer to the Martens clause seem symptomatic of the current state of affairs. In the absence of other adequate instruments of international law which combine legal and moral norms (Sparrow, 2017), the Martens clause appears as a convenient negotiation tool, the evocation of which can resolve potential regulatory impasses. However, this usefulness of the Martens clause can also be illusory due to the well-known difficulties and discrepancies in its interpretation, in particular the lack of a uniform agreement regarding the notions of “the laws of humanity and the dictates of public conscience.” Nevertheless it is worth noting Judge Shahabuddeen’s views on the functionality of the Martens Clause. In his dissenting opinion to the Advisory opinion in the case *Legality of the Threat or Use of Nuclear Weapons* (1996), he stressed that in order to determine the meaning of the “laws of humanity and dictates of public conscience”, they must be considered in the light of evolving circumstances, including technological changes and the development of means and methods of warfare, and the outlook and the tolerance levels of the international community.

The requirement to update the concepts mentioned above in the case of the development of disruptive military technologies perfectly fits into the discussion on the regulatory difficulties related to LAWS, especially because of the flagged ethical challenges that require going beyond the binding legal regulations. In this sense, the Martens clause expresses a sense of justice and morals as sources of law. An in-depth debate and reflection on the objectives, nature, and manner of the modern conduct of hostilities could provide an essential foundation for the subsequent regulation of disruptive technologies.

In this regard, the development of LAWS should be considered in the context of the principles of humanity and the dictates of public conscience. While the former should be understood broadly as respect for human dignity and moral standards shared by humanity, the latter could add a fresh perspective by referring to relevant public opinion. Since the general public opinion reflects the current state of the international climate, it seems advisable to use civil society (e.g. NGOs) or experts in various fields of social science (e.g. doctrinal positions expressed in military manuals and scientific publications) to properly determine international morality. Public conscience is a kind of opening of IHL to non-legal factors. Veuthey rightly points out that:

> Public conscience is not the monopoly of legal experts. More research is needed in history, anthropology, spirituality, international relations, in order to defend fundamental rights pertaining to life and human dignity in all situations (Veuthey, 2004, p. 631).
In addition, it is worth noting that although the perception of public conscience through the prism of opinion polls may seem too narrow an approach, it is thanks to modern technological possibilities that collecting large amounts of feedback from societies can be a starting point for a discussion about the state of public morality (Roff, 2017). Qualified public opinion can be also conceived as „a host of draft rules, declarations, resolutions, and other communications expressed by persons and institutions highly qualified to assess the laws of war although having no governmental affiliations” (Written Statement of the Government of Nauru, 1994, p. 68).

There is also a second understanding of the ‘dictates of public conscience’ that equates them with the ethos of our time – human rights law (Milanović, 2019). But as Kałduński (2009) rightly points out, reducing the importance of the dictates of public conscience and principles of humanity to a simple reference to human rights standards would deprive the Martens clause of its character as a self-sufficient and independent legal standard.

Still, consideration of both spectrums in the debate on the regulation of disruptive technologies, and especially LAWS, could be of great benefit and would definitely contribute to advanced inclusive politics. While assuming an awareness of the ineffectiveness of the Martens Clause on the part of experts and representatives of States, it should be noted that the international community engaged in discussions about LAWS seeks a construction that will provide answers to questions that go beyond mere legal analysis and safeguard the interest of mankind (including its survival) against the „whims of States and people's weaknesses” (Veuthey, 2004, p. 635). Taking into account the theoretical difficulties and recognizing the need to concretize the meaning of the Martens clause, it should be recognized that the discussion on LAWS is another convenient moment (after missed opportunities on the occasion of regulation of laser weapons or landmines) to define the meaning and function of the Martens clause in the processes leading to the adoption of new international legal obligations. Because of its incorporation of moral norms, the Martens clause has the potential to more adequately adapt international law standards to the expectations of the contemporary international society.

**Ethics as a Part of the Decision-Making Process**

Ethical considerations should be incorporated into the execution of State power on at least two levels: the examination of the legality of the means and methods of warfare accordingly to Art. 36 API; and in the targeting process.

In accordance to the ICRC’s Guidelines, the Martens clause shall be included in the examination of the legality of the specific means of warfare in question as one of the rules of international law in force for a given State (ICRC, 2006). Therefore, the State should recognize as illegal all means and methods of warfare that are incompatible with principles of humanity and the dictates of public conscience. In the course of research into the mechanisms of legal review of new weapons systems, it was pointed out that one of the practical implications of
this standard is the need to send enquiries to the relevant government structures representing political authority (and not to members of armed forces or legal advisers) to determine whether a given developed or acquired weapon does not contradict the principles of humanity and the dictates of public conscience on a strategic level. This mechanism assumes that a State has an armament strategy not susceptible to policy changes, and that policy changes if they occur reflect the will of the nation (this argument holds in the case of democratic regimes, but not in rogue States). The ability to identify such public opinion would permit the issuance of an opinion on the legality of a particular reviewed weapon in accordance with the prevailing military strategy, including in cases of moral decisions regarding the type of equipment developed or acquired to defend one's own citizens (German defence minister backs European armed drone, 2014). Making use of an appeal to the public and leaving to them a decision regarding the compliance of a particular weapon with the Martens clause would solve the dilemma of understanding and assessing non-legal standards by the team or persons responsible for the legal review of the weapon.

Seen in this light, despite the controversy at the theoretical level the Martens clause has the normative potential to have a direct impact on political decisions, whether at the international or national level. In particular, the clarification of the prevailing moral and legal standards set by it can affect the improvement of the implementation and enforcement of IHL. This is all the more important in the case of the regulation of LAWS, as the main challenges relate to extra-legal questions, in particular whether we should transfer the decision-making process (regarding killing people) to artificial intelligence and what level of uncertainty we are ready to accept. These complex issues should be addressed first and foremost from the ethical perspective. It involves such fundamental questions as the place of humans in targeting processes, the understanding of human dignity during war, as well as the nature and objectives of modern armed conflicts. Since IHL norms do not refer explicitly and conclusively to these matters, the international community should „take a step back” and contemplate the ethical implications of disruptive technologies, and only after such a contemplation adopt the relevant legal norms.

In this spirit, it should be noted that part of the doctrine has undertaken research which has led it to adopt new concepts based on the analogy between the targeting process and other legally regulated situations.

Lieblich and Benvenisti (2016) proposed an original perspective which combines human rights with the administrative law perspective. The authors suggest a re-evaluation of the debate on LAWS and to perceive the targeting process – using the theoretical bridge of the procedural justice and due process paradigms – as a sphere of the execution of administrative power, understood as the exercise of State power over persons through a computerized proxy.

The authors refer to the complementarity of norms belonging to different legal regimes, and in this case especially to second-order norms, i.e. the concepts of transparency and discretion, as basic principles of administrative law relating to the procedural aspect of
decision-making including, in the context of armed conflicts, targeting (Lieblich, Benvenisti, 2016). The authors present the praiseworthy view that it is justifiable to treat contemporary armed conflicts as a form of exercise of State power over human beings, rather than as a relationship between two sovereign entities. Therefore, once again the issue of human dignity and adequate treatment of human beings arises. However, this adequacy remains to be defined in the context of the robotization of all procedures, starting with creditworthiness and sentencing and ending with killing (this issue centers around the scope of the notion of „meaningful human control”). Last but not least, the sovereignty of the State is pictured as „a trusteeship of humanity”, which means a commitment to take into account the perspective of „others” (for example civilians and combatants of the opposing side) in the decision-making processes and to bear responsibility for the decisions taken (Benvenisti, 2013).

The above mentioned approach aggregates the arguments (norms) from different legal regimes that should be exploited by States in the process of lawmaking, while reflecting on the state of contemporary international law. This should allow them to better deal with the key challenges created by LAWS. Such an intersectional method consolidates the systemic character of international law and better exposes the standards and needs of humankind in the context of the advent of robotized warfare.

**Conclusions**

Having regard to the role of the Martens clause as an opening to incorporate ethical considerations and conclusions stemming from other legal regimes, I argue that IHL is gradually yet incessantly striving for further humanization, understood as prioritization of the value of human life and role of the human factor (both its presence and behavior) in decision-making procedures relating to the health and life of victims of modern armed conflicts.

Nevertheless, humanity and military necessity, the two fundamental considerations of IHL – which are supposedly incorporated in the basic principles of IHL both practically and theoretically – need to be balanced as a result of negotiations and the law-making process. Therefore, it may be argued that the time has now come for States to make decisions with respect to the qualitative shift in the nature of armed conflicts, i.e. the transfer of the decision-making process from human to artificial intelligence. As this is being done in the civil sphere, all the more so should it be expected in the military sphere. The place of artificial intelligence among human operators has to be defined by States, and in accordance with the Martens clause they need to adopt a modern understanding of humanity and the dictates of public conscience.
References


Author

Kaja Kowalczewska
PhD, kowalczewska.kaja@gmail.com - Pedagogical University of Krakow, Institute of Law, Administration and Economics