Cyber Operations and The Humanization of International Humanitarian Law: Problems and Prospects

Giacomo Biggio*

The aim of International Humanitarian Law ("IHL") is to regulate the conduct of hostilities while, at the same time, balancing the two overarching concepts of military necessity and humanity. While the principle of military necessity allows a party to a conflict to exercise any amount of armed violence which is necessary for the accomplishment of a military purpose, the principle of humanity aims at minimizing the amount of *physical violence* caused to combatants and the civilian population. From the late 19th century onwards the principle of humanity has progressively eroded the domain of military necessity, influencing the creation and interpretation of IHL, in a process which is referred to as the "humanization" of IHL[1]. A key area of IHL in which such process has taken place is the law of targeting, whose aim is to limit "attacks"[2], by prohibiting belligerents to direct them against civilians. In this context, the rise of cyber warfare capabilities establishes a tension between the violence-centered *rationale* of the law of targeting and the nature of cyber attacks, as their effects may have devastating consequences even without causing any form of *physical violence*.[3] What kind of cyber operations should qualify as 'attacks' under the law of targeting? The answer to this question can reconfigure the delicate balance between military necessity and humanity, raising implications for the protection of the civilian population and the humanitarian aims of IHL.

My contribution offers a critical evaluation of the relationship between the principle of humanity within the law of targeting and cyber warfare, and proposes a "human security" paradigm for the regulation of cyber attacks in armed conflict.

I. INTRODUCTION

The term "humanization of warfare"[4] denotes a process in which humanity-oriented considerations influence the creation and interpretation of the rules of

---

* PhD Candidate, University of Sheffield.

1 Meron, T., "The Humanization of Humanitarian Law" (2000) 94 Am J Intl L 239-278.

2 An attack is "an act of violence against the adversary, whether in offer or defense": Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protections of Victims of International Armed Conflicts (Protocol I), 8 June 1977, Art. 49 (1).

3 Consider, for instance, a cyber operation that targets the stock market of a state or one that shuts down an electrical power grid of a city.

4 Meron, *supra* note 1. See also Kjetil Mujezinovi Larsen, Camilla Guldahl Cooper & Gro

International Humanitarian Law (IHL). This process, which is the reflection of a trend that embraces the whole spectrum of international law,[5] manifests itself in different areas of IHL, from the rules aimed at improving the conditions of Prisoners of War, medical and religious personnel deployed in the battlefield to those designed to enhance the protection of the civilian population and civilian objects. With regard to the latter aspect, humanitarian considerations play a fundamental role in the law of targeting, a segment of the *jus in bello* consisting in the rules of distinction, proportionality and precaution, whose primary purpose is to regulate "attacks", a concept defined under Art. 49 of the First Additional Protocol to the Geneva Conventions (AP I) as "acts of violence against the adversary, whether in offence or in defence."[6] Once an act qualifies as an attack, the principle of distinction requires combatants to direct their attacks against military objectives, and forbids attacks against civilians[7] — individuals who are not directly participating in hostilities by performing or participating in the planning of attacks[8]; the principle of proportionality prohibits indiscriminate attacks, that is, attacks that cause incidental loss of life or injury to civilians, damage or destruction to objects, or a combination thereof, which would be excessive in relation to the military advantage expected from the operation.[9] Lastly, the principle of precaution requires belligerents to take a series of precautionary measures when launching attacks, such as verifying that a target of an attack is a military objective,[10] suspending or cancelling the attack when the target is a civilian object or when the attack would be indiscriminate,[11] and taking all feasible precaution in their choice of means and methods of warfare in order to prevent and minimize the adverse effects of attacks on the civilian population.[12] The rationale that underlies the law of targeting is, therefore, to limit the amount of violence that the parties to a conflict may lawfully employ during an armed conflict in order to give civilians the greatest possible degree of

---

Nytsuen, eds, *Searching for a "Principle of Humanity" in International Humanitarian Law* (New York: Oxford University Press, 2013).

5 Ruti G. Teitel, *Humanity Law* (New York: Oxford University Press, 2011); Anne Peters, "Humanity as the A and of Sovereignty" (2009) 20:3 EJIL 545.

6 *Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts (Protocol I)*, 18 June 1977, 1125 UNTS 3 art 49 (entered into force 7 December 1978) [AP I].

7 *Ibid* art. 48.

8 *Geneva Convention Relative to the Treatment of Prisoners of War*, 12 August 1949, 75 UNTS 135, art. 3 (entered into force 21 October 1950). See also Nils Melzer, *Interpretive Guidance on the Notion of Direct Participation in Hostilities Under International Humanitarian Law* (May 2009), online: International Committee of the Red Cross <www.icrc.org/eng/assets/files/other/icrc-002-0990.pdf>.

9 AP I, *supra* note 6, art. 51(5)(b).

10 *Ibid* art. 57(2)(a)(i).

11 *Ibid* art. 57(2)(b).

12 *Ibid* art. 58.

protection. While the law of targeting represents a defining moment in the humanization of IHL, the rise of cyber warfare and the use of Computer Network Attacks (CNAs)[13] in the midst of hostilities may adversely affect such process. In fact, the notion of attack — the kernel of the law of targeting — is premised on the concept of violence, which has been interpreted as a synonym of *physical harm*:[14] everything that causes injury or death to civilians, or damage or destruction of objects as its primary effect qualifies as an attack, whereas acts not causing physical violence are not limited by the rules on distinction, proportionality and precaution. Beyond physical harm, it must be added that the notion of violence under AP I includes military harm as well. As correctly observed by the ICRC, the notion of military harm "should be interpreted as encompassing not only the infliction of death, injury, or destruction o[f] military personnel and objects, but essentially any consequence adversely affecting the military operations or military capacity of a party to a conflict", such as acts of "sabotage and other armed or unarmed activities restricting or disturbing deployments, logistics and communications."[15]

CNAs are, on the one hand, clearly capable of causing violent consequences: incidents such as the Stuxnet worm attack, which targeted the centrifuges of the Iranian nuclear facility of Natanz, causing the destruction of physical components of the nuclear plant and the release of radioactive materials in the surrounding environment, demonstrate how the causation of physical violence is well within the reach of cyber warfare. The same holds true for the causation of military harm, as a computer network attack can interfere with the Supervisory Control and Data Acquisition System (SCADA system) of any modern weapon system.

To this extent, the humanity-inspired rules on targeting can be easily interpreted to govern cyber attacks taking place in an armed conflict. On the other hand, CNAs may cause no physical violence at all, yet can have serious consequences: the paradigmatic example is a concerted cyber attack against the banking system of a State, aimed at causing a financial collapse. It is no doubt that, even if no death or destruction would occur, the target State would be on the brink of collapse: the consequences of such an attack, as it will be demonstrated in the course of the article, represent a new form of violence, even if non-physical in nature.

---

[13] A cyber attack or Computer Network Attack is defined by the US Joints Chief of Staff as an operation designed to "disrupt, deny, degrade, or destroy information resident in the target information system or networks, or the systems/ networks themselves": James E. Cartwright, *Memorandum for Chiefs of the Military Services Commanders, Commanders of the Combatant Commands, Directors of the Joint Staff Directorates,* at 3, online: <www.nsci-va.org/cyberreferencelib/2010-11-joint%20terminology%20for%20cyberspace%20operations.pdf>.

[14] Michael N. Schmitt, "Wired Warfare: Computer Network Attack and *Jus in Bello*" (2002) 84:846 Intl Rev Red Cross 365 at 377.

[15] Melzer, *supra* note 8 at 47-48.

The extent to which the notion of "attack" should be applied to non-physical violent cyber operations can dramatically impact upon the humanization of warfare. Should a non-physical violent CNA be considered as not falling within the definition of "attack" under Art. 49 AP I, then the law of targeting would not apply, and the parties to the conflict would be free to launch this kind cyber operations without legal scrutiny.

The purpose of this article is to present a convincing case for the application of the rules of targeting to non-physical violent CNAs. In order to do so, the article will address and critique the main interpretive approaches to the notion of cyber attack in the *jus in bello* regime, then propose an alternative approach, based on human-security oriented concerns, to the notion of cyber attack.

II. THE *TALLINN MANUAL* INTERPRETATION OF CYBER ATTACK: PROBLEMS AND CRITIQUES

The most authoritative effort to apply Art. 49 AP I to cyber attacks comes from the *Tallinn Manual on the International Law Applicable to Cyber Warfare* ("Tallinn Manual"), which represents the collective effort of a Group of Experst (GoE) convened on behalf of NATO with the task of clarifying the applicability of the *jus in bello* and the *jus ad bellum* to cyber warfare. A cyber attack is defined, under rule 30 of the Tallinn Manual, as a "cyber operation, whether offensive or defensive, that is reasonably expected to cause injury or death to persons or damage or destruction to objects".[16]

The Tallinn Manual's approach analogizes the effects of cyber attacks to those caused by kinetic weapons, in what has been referred to as the "kinetic equivalence effects test" (KEE Test),[17] to the extent that a cyber attack causes physical violence in the form of injury to individuals or damage to objects, it is an attack for the purposes of IHL. Conversely, the accompanying commentary notes that "non-violent operations, such as psychological cyber operations or cyber espionage, do not qualify" as such.[18]

On the one hand, the KEE Test has the merit of being easily applicable to some clear-cut cases of highly-destructives cyber attacks. On the other hand, its limits appear evident when it comes to classifying the nature of cyber operations that do not fall into higher extremity of the causation of physical violence (such as cyber attack causing a plane to crash) or at its opposite (as in the case of an operation of data exfiltration). Secondly, under the KEE Test, the classification of cyber operations that interfere with the functionality of an object remains

---

[16] Michael N. Schmitt, ed., *Tallinn Manual on the International Law Applicable to Cyber Warfare* (New York: Cambridge University Press, 2013) at 106 [Schmitt, *Tallinn Manual*].

[17] Karine Bannelier-Christakis, "Is the principle of distinction still relevant in cyberwarfare?" in Nocholas Tsagourias and Russell Buchan, eds., *Research Handbook on International Law and Cyberspace* (Edward Elgar Publishing Limited, 2015) 343 at 348.

[18] Schmitt, *Tallinn Manual*, *supra* note 16 at 106.

unclear. Consider, for instance, a CNA that shuts down the electrical power grid of a city. Albeit the Tallinn Manual excludes the possibility that such an operation may fall under the meaning of "cyber attack", it has specified that the threshold for the qualification requires "physical replacement of the components" of the systems whose functionality has been impaired or compromised. Should the cyber operation on the electrical power grid require physical replacement of some of its components, that alone should qualify the cyber operation as an attack.

Vice versa, a cyber operation that requires mere digital data restoration does not qualify as an attack according to the majority of the GoE,[19] even in the presence of other grave consequences (such as massive financial losses). Under the Tallinn Manual definition, these operations would not qualify as attacks as the threshold of physical violence would not be met. The reason behind this approach lies in the fact that the Tallinn Manual interprets digital data as an immaterial entity, something which differs from an "object" because it is not "visible and tangible", as noted by the ICRC Commentary on the Additional Protocols.[20] It follows that, if data is not an object, then a cyber attack that merely alters or deletes digital data does not cause "damage" or "destruction" to objects, and no physical violence occurs. This leads to a further consideration: what if a cyber attack interferes with digital data stored in the system of a military objective (for instance, an operation that disables the SCADA system of anti-aircraft artillery) in a way that does not require physical replacement? In such a case, it could well be argued that the operation would adversely affect the military capacity of one of the parties to an armed conflict. However, under the Tallinn Manual approach, it is debatable whether such an operation would qualify as an attack.

Against the Tallinn Manual interpretation several counter-arguments have been put forward. Nils Melzer suggests that what is relevant for the application of the rules of distinction, proportionality and precaution is the concept of "military operation" and not the notion of "attack".[21] Military operations are defined by Art. 48 and 51 of AP I as "all movement and acts related to hostilities that are undertaken by armed forces."[22] Hence, even a cyber operation that causes non-physical violent consequences and would not, under the Tallinn Manual interpretation, qualify as an attack, would nonetheless fall within the broader notion of military operations and would have to comply with the law of targeting. While this interpretation has the merit of circumventing the

---

19 *Ibid* at 108-109.

20 Yves Sandoz, Christophe Swinarski & Bruno Zimmerman, eds., *Commentary to the Additional Protocols of 8 June 1977 to the Geneva Conventions of 12 August 1949* (Geneva: International Committee of the Red Cross, 1987) at paras 2007-2008.

21 Nils Melzer, *Cyberwarfare and International Law* (2011) at 27 online: United Nations Institute for Disarmament Research < unidir.org/files/publications/pdfs/cyberwarfare-and-international-law-382.pdf > .

22 Sandoz, *supra* note 20 at para 1875.

interpretive problems associated with the definition of "attack" in the cyber context, it does not stand against a systematic reading of Part IV of AP I, which deals with the protection of civilians and civilian objects: as Roscini points out, the ICRC commentary clarifies that section IV applies only to attacks, that is, to "military operations during which violence is used."[23] Furthermore, AP I operates a distinction between attacks and military operations with regards to the obligation that belligerents have to comply with in order to protect the civilian population and objects: while attacks are limited by the principle of distinction, military operations are merely subject to the more generic obligation to take constant care. The proposed interpretation is therefore untenable.

A different view has been put forward by Dormann, who argues that what makes an act an "attack" does not include mere destruction of objects, but also other modalities of interaction, namely capture and neutralization, as suggested by the definition of "military objective" under Art 52 (2) of AP I.[24] More specifically, the definition of "neutralization" appears to suit the reality of cyber operations, since it was included to signify "an attack for the purpose of denying the use of an object to the enemy without necessarily destroying it."[25] The main weakness of this approach is that it relies on the notion of military objective to define what is an "attack". However, the definition of military objective implies the existence of an attack in the first place, a notion which is already defined in the Protocol and which, as it has been discussed above, is dependent upon the causation of physical violence. On the other hand, the concepts of "capture and neutralization" do not operate as qualifiers of an attack: instead, they are just descriptions of what is the objective of an attack launched against a military objective, in the sense that an attack can only cause not only the destruction of a military objective, but also to its capture or neutralization. The interpretation but forward by Dormann, being limited to an examination of military objectives, can be traced back the rationale that a cyber operation causing military harm — that is, affecting the military capacity of the adversary — by neutralizing an object would constitute an attack. However, it does not explain whether cyber operations directed against civilian objects, and which do not result in physical violent consequences, should be qualified as attacks or not.

---

[23] Marco Roscini, *Cyber Operations and the Use of Force in International Law* (Oxford: Oxford University Press, 2014) at 178.

[24] Knut Dormann, *Applicability of the Additional Protocols to Computer Network Attacks* (2004) at 6, online: International Committee of the Red Cross <https://www.icrc.org/eng/assets/files/other/applicabilityofihltocna.pdf>.

[25] Michael Bothe, Karl Josel Partsch & Waldemar A. Solf, *New Rules for Victims of Armed Conflicts* (Martinus Nijhoff Publishers, 1982) at 325.

III. THE EVOLUTIVE INTERPRETATION OF THE NOTION OF "VIOLENCE" WITHIN THE DEFINITION OF "ATTACK"

While the former approaches focused on aspects external to the notion of attack, namely the definition of military operation and the modalities of interaction with military objectives, several other authors have proposed an "evolutionary" or "dynamic' interpretation of the notion of attack and the underlying concept of violence.[26] Evolutionary interpretation is a tool used to clarify the meaning of terms within a treaty, and it is premised on the assumption that the ordinary meaning of a certain term can evolve over time. An explanation of this interpretive technique is found in the Navigation Rights case before the ICJ, where the court held that if the parties choose a generic term in a treaty entered into force for a very long time, they should be presumed to have intended for such a term an evolving meaning.[27]

While the Navigations Rights case proves its usefulness in explaining the evolutionary interpretation principle, other pronouncements by the ICJ offer practical examples of the application of such approach. The most relevant one is to be found in the ICJ's Advisory Opinion on the Legality of the Threat or Use of Nuclear Weapons. Faced with the question whether, absent any express treaty provision that banned the use of nuclear weapons, their use should have been considered lawful under IHL, the court argued that excluding nuclear weapons for the application of the principles of IHL "would be incompatible with the intrinsically humanitarian character of the legal principles in question which permeates the entire law of armed conflict and applies to [. . .] all kinds of weapons, those of the past, those of the present and those of the future."[28] The argument of the court, while related to the more generic issue of the applicability of IHL to a certain weapon technology, is however an unquestionable example of evolutive interpretation of the law. As a further example, in the Gabcikovo-Nagymaros case, the ICJ considered that the terms of the bilateral agreement between Czechoslovakia and Hungary should be interpreted taking into consideration the law and scientific knowledge as they evolved at the time the case was decided before the court and not at the time the treaty was concluded.[29] A dynamic approach has been, finally, endorsed by other international courts, such as the European and the Inter-American Court of Human Rights, which interpret human rights treaties as "living instruments".[30]

---

[26] Bannerlier-Christakis, *supra* note 17 at 354-355; Kubo Maák, "Military Objectives 2.0: The Case for Interpreting Computer Data as Objects under International Humanitarian Law" (2015) 48:1 Israel LR 55 at 68.

[27] *Dispute Regarding Navigational and Related Rights (Costa Rica v Nicaragua)*, [2009] ICJ Rep 213 at para 66.

[28] *Legality of the Threat and Use of Nuclear Weapons*, Advisory Opinion [1996] ICJ Rep 226 at para 86.

[29] *Gabcikovo-Nagymaros Project (Hungary v Slovakia)*, [1997] ICJ Rep 7 at para.140.

[30] Maák, *supra* note 26 at 73.

Assuming that the evolutionary interpretation represents a useful tool for understanding the notion of violence in the present day, the main question then revolves around what elements, within the notion of violence, need to be re-interpreted.

According to this strand of the doctrine, the main issue revolves around the interpretation given by the Tallinn Manual to digital data as an immaterial entity.[31] Thus, Maák points out that digital data may, in fact, qualify as an "object" under IHL, since the drafters of the Protocols interpreted objects as being "tangible and visible" in opposition to abstract concepts, such as "the general objective [. . .] of a military operation."[32] The reason behind such distinction is apparent if one considers that "[i]f a party's aim amounted to a legitimate target justifying an attack by its opponent, the detailed and balanced rules on targeting would lose any sense."[33] Following this logic, if digital data is an object, an act that results in damage or destruction with digital data would amount to an attack.

In light of these considerations, the view according to which cyber operations that target digital data fall within the same category as psychological operations cannot be shared. As Lubell suggests, the aim of psychological operation consists in persuading a certain target, be it the military or the civilian population, by influencing its morale. On the other hand, cyber attacks are "more often designed with some form of harmful effect in mind [. . .] even if not always measurable in casualties."[34]

In conclusion, the dynamic approach focuses primarily on the inclusion, within the concept of "object", of digital data. If digital data is an object that can be attacked, then it follows that interference with digital data, either in the form of destruction or alteration, falls within the meaning of physical violence. According to this view, a cyber operation that interferes with the functionality of a military objective by targeting its digital data shall be considered an attack, since it adversely affects the military capacity of the belligerent by causing military harm. While the inclusion of this kind of cyber operations within Art. 49 AP I is a step forward compared to the Tallinn Manual's approach, some authors have pointed out that the dynamic interpretation of the notion of attack is too extensive.[35] In fact, if a cyber attack is an operation that deletes, alters or in any other way interferes with digital data, then any cyber operation would qualify as an attack under IHL, be it an operation of data exfiltration from a governmental website or a cyber attack aimed at shutting down an electrical power grid. Moreover, participation in hostilities would be determined by the

---

[31] *Ibid* at 70; Noam Lubell, "Lawful Targets in Cyber Operations: Does the Principle of Distinction Apply?" (2013) 89 Intl L Studies 252 at 267.

[32] Maák, *supra* note 26 at 60.

[33] *Ibid*.

[34] Lubell, *supra* note 31 at 263-264.

[35] Cordula Droege, "Get off My Cloud" (2012) 94:886 Intl Rev Red Cross 533 at 558-559.

performance of any of such acts. Such a legal outcome, therefore, cannot be agreed upon because it would put too many restraints on the conduct of hostilities and it would likely be disregarded by the parties to a conflict, leading to non-compliance with the rules of targeting and would have an adverse impact on the humanization of IHL. For these reasons, the dynamic approach does not represent a viable alternative to the Tallinn Manual's interpretation.

IV. THE RELATIONSHIP BETWEEN EFFECTS OF CYBER ATTACKS AND THE NOTION OF *VIOLENCE*: A HUMAN SECURITY-BASED PARADIGM

A proper humanity-oriented interpretation of the notion of attack should focus on the relationship between the nature of the effects caused by cyber attacks and the rationale behind the qualification of an act as an attack.

Conventional attacks, as well as cyber attacks, can have different effects in terms of their causal proximity to the act that has produced them. In this regard, a distinction must be made between first and second-order effects. First order effects are those which are directly consequential to an attack. In the case of a conventional attack, the first order consequences always result in the causation of death or injury to individuals or destruction of objects, as in the case of an air strike that destroys an ammunition factory. With regards to cyber attacks, their first order effects always result in the interaction between the cyber operation and its target, that is, digital data stored in a computer system or a network. As an example, the first-order effects of a cyber operation directed against an electrical power grid stem from the interaction between the cyber attack and the digital data within the operating system of the electrical grid. In this regard, a cyber operation that interferes with digital data may qualify as an attack only when it is directed against a military objective, since it would adversely affect the military capacity of a belligerent by inflicting military harm. However, when such a cyber operation is directed towards a civilian object, mere destruction or interference with digital data alone does not result in the infliction of physical violence, and therefore is not sufficient for the qualification of a cyber operation as an attack. What qualifies a cyber operation as an attack is the evaluation of its second-order effects, which can be described as the consequences of the first-order effects. Second-order effects can result in physical violence: for instance, considering the above example, the loss of electricity resulting from shutting down the power grid may cause a fire which can cause death or injury to civilians, or damage or destruction to civilian objects. Moreover, non-physical second-order consequences may also take place, as shutting down a grid which provides electricity for a town would deprive the civilian population of a basic service and would likely generate panic. As already noted, the Tallinn Manual recognizes that, as long as second-order effects cause physical violence, then the cyber operation that produced them is an attack, as evidenced by rule 30, according to which:

> The word "cause" [. . .] is not limited to effects on the targeted cyber system. Rather, it encompasses any reasonably foreseeable consequential damage, destruction, injury, or death. Cyber-attacks seldom involve the release of direct physical force against the targeted cyber system; yet, they can result in great harm to individuals or objects. For example, the release of dam waters by manipulating a SCADA system would cause massive downstream destruction without damaging the system. Were this operation to be conducted using kinetic means, like bombing the dam, there is no question that it would be regarded as an attack. No rationale exists for arriving at a different conclusion in the cyber context.[36]

Therefore, second-order non-physical consequences are not considered violence under the Tallinn Manual approach, regardless of their adverse effects on the target State or the civilian population.

In this regard, it is submitted that there should be an interpretive shift in the rationale behind the prohibition on attacks, which is premised on the protection of State Security oriented interests, with values influenced by the notion of Human Security. The concept of "human security", first introduced in the UN Human Development Report 1994,[37] is an emerging practice that aims at integrating the traditional notion of State Security, focused on the protection of the territorial integrity of a State against external aggression, with an approach focused on securing of essential human needs, such as environmental security, economic security, health and food security and political security among others.[38] The adoption of a human security based paradigm for the definition of cyber attacks under the *jus in bello* would allow to conceive the notion of violence as including not only second-order physical violent consequences, but also second-order non-physical consequences that have an adverse impact on Human Security protected needs. The focus on basic human needs is an element which is also present in several specific provisions of AP I that take into account second-order consequences that are different from physical violence. For instance, Art. 51(2) prohibits attacks aimed at spreading terror among the civilian population:[39] the notion of terror denotes a form of severe psychological suffering that is certainly different from the notion of "injury". Similarly, Art. 55 protects the natural environment against "widespread, long-term and severe damage"[40]: while the notion of "damage" includes destruction of parts of the

---

[36] Schmitt, *Tallinn Manual*, *supra* note 16 at 107.

[37] *Human Development Report 1994* (Oxford: Oxford University Press, 1994), online: United Nations Development Programme >.

[38] Martin Wählisch, "Human Security: Concept and Evolution in the United Nations" in Frauke Lachenmann, Tilmann J. R—der & Rüdiger Wolfrum, eds, *Max Planck Yearbook of United Nations Law* (Boston: Brill Nijhoff, 2014) 3 at 18; Matthew S. Weinert, "From State Security to Human Security?" in Patrick Hayden, ed., *The Ashgate Research Companion to Ethics and International Relations* (Routledge, 2009) 151.

[39] AP I, *supra* note 6, art 51 (2).

environment, it can also encompass other forms of alteration of the natural environment which cause environmental harm, such as massive oil or radioactive materials spills. Finally, Art. 54 prohibits to "attack, destroy, remove or render useless objects indispensable to the survival of the civilian population"[41]: in this case, the damage suffered from the civilian population from an attack targeting "objects indispensable" to their survival cannot be understood as being equivalent to "death" or "injury" but, rather, is more akin to a form of deprivation that severely affects its health and its well-being. It can be argued that IHL is concerned with the regulation of second-order effects where they do not appear to cause physical harm in the form of death or injury to individuals or material damage or destruction to objects. The reason of this statement lies in the fact that the second-order consequences listed under Arts. 51(2), 55 and 56 AP I are capable to endanger values protected by IHL and which are, at the same time, falling under the wider concept of human security, such as the safety of environment, the health of the civilian population and its psychological well-being: in these circumstances, these second-order effects can be qualified as a form of non-physical violence. Against these considerations, it can be noted that what triggers the application of these provisions is the existence of an attack that causes first-order physical violent consequences. Therefore, in the case of a CNA that would not cause any death or injury to individuals, or damage or destruction to objects, Articles 51 (2), 54 and 56 would have no reason to apply. However, it is unreasonable to exclude from the definition of attack a cyber operation whose first-order effects consist in mere interference with digital data, but whose second-order effects would cause non-physical violent consequences that would adversely affect the security of a state in the same way as physical violent consequences would. Rather, such a cyber operation should be regarded as an "attack" under IHL: under this approach, a cyber operation aimed at spreading terror against the civilian population would fall under the meaning of Art. 51(2), whereas one that deactivates the control system of an oil plant resulting in massive oil spills would be prohibited by Art. 56.

The above interpretation expands the notion of violence to non-physical consequences that, by endangering values that are protected by IHL, adversely affect the security of humans. This consideration begs the question of whether cyber operations whose consequences would severely affect human security-related values which are not explicitly protected by the provisions of AP I should be considered new forms of "attacks" under IHL. The answer is in the affirmative.

Consider again the scenario in which a CNA targets the stock market of the victim state by interfering with the digital data stored within the system operating the stock market, the CNA would cause second-order consequences resulting in massive currency manipulation leading to deflation, and the freezing of hundreds

---

[40] *Ibid* art 55.

[41] *Ibid* art 54.

of thousands of bank accounts. The kind of consequences will not be physically violent in nature, however their outcome would result in widespread unemployment, severe distress among the civilian population, and would be considered as a threat to the economic and financial stability of the target State. Under this scenario, the second-order effects of the attack would be extremely serious, causing economic harm by depriving the civilian population of the possibility to withdraw currency. This kind of consequences should be qualified, as they endanger human security-related values, as a form of non-physical violence and the cyber operation that caused them should be considered as an attack within the meaning of Art. 49 AP I.

In order to determine with greater certainty whether a cyber operation is intended to cause second non-physical consequences that may threaten essential human needs, it is useful to examine the notion of Critical National Infrastructure (''CNI''). Albeit there is no accepted definition of CNI under international law, different states have defined the term in a similar manner. For instance, the US Joint Terminology for Cyberspace defines them as ''[s]ystems and assets, whether physical or virtual, so vital that the incapacity or destruction of such may have a debilitating impact on security, economy, public health or safety, environment, or any combination of these matters, across any Federal, State, regional, territorial, or local jurisdiction.''[42] The German definition is also similarly worded, since it includes ''organizations or institutions with major importance for the public good, whose failure or damage would lead to sustainable supply bottlenecks, considerable disturbance of public security or other dramatic consequences.''[43] The European Commission also defines the notion of European Critical Infrastructure as ''an asset, system or part thereof located in Member States which is essential for the maintenance of vital societal functions, health, safety, security, economic or social well-being of people, and the disruption or destruction of which would have a significant impact in a Member State as a result of the failure to maintain those functions.''[44]

In this regard, it must be noted that the definitions of CNI, whether at State or regional level, represent a perfect combination between the notion of State Security and Human Security, since the destruction or disruption of the functions of a CNI would affect several services of vital importance such as energy, food, water, transportation, banking, communication, financial and governmental sectors.[45] It can therefore be assumed that a cyber operation that disables,

---

[42] Cartwright, *supra* note 13 at 5.

[43] *Cyber Security Strategy for Germany,* (February 2011) at 15, online: Federal Office for Information Security <https://www.bsi.bund.de/SharedDocs/Downloads/EN/BSI/Publications/CyberSecurity/Cyber_Security_Strategy_for_Germany.pdf>.

[44] EC, *Council Directive 2008/114/EC of December 8 2008 on the identification and designation of European critical infrastructures and the assessment of the need to improve their protection*, [2008] OJ, L 345/75 art 2(a).

[45] See e.g. *The UK Cyber Security Strategy: Protecting and Promoting the UK in a Digital*

temporarily or permanently, the functionality of a CNI would cause consequences that would adversely affect Human Security related needs.

For these reasons, it is submitted that the definition of attack in the cyber context should be interpreted as including the following categories. Firstly, there are cyber operations whose second-order consequences are intended to cause death, injury to civilians, destruction or damage to objects; these cyber operations qualify as attacks because they cause physical violence. Secondly, cyber operations which cause mere interference with digital data would also qualify as an attack if they are intended to adversely affect the military capacity of one of the belligerents, thereby causing military harm.

The last category includes cyber operations that cause non-physical violence, namely, cyber operations that are intended to cause psychological harm, environmental harm, those which deprive the civilian population of objects necessary for its survival, and any other cyber operation which adversely impacts human-security protected values, such as CNAs that are intended to interfere with the functionality of a CNI.

In conclusion, this article has shown how different interpretations of the concept of cyber attack may have an impact on the humanization of IHL. Discussing the prevalent approaches in the doctrine, it demonstrated how they would adversely affect the humanization process. It has, then, proposed an alternative view that consisted on incorporating human security related considerations in the concept of violence, upon which the definition of attack under Art. 49 AP I is based. By doing so, the principle of humanity can continue to play an important role in the cyber scenario, requiring belligerents to comply with the rules on distinction, proportionality and precaution when launching CNAs, and, therefore, increasing the protection of the civilian population from their effects.

World, (November 2011) at 9, online: < www.gov.uk/government/uploads/system/uploads/attachment_data/file/60961/uk-cyber-security-strategy-final.pdf > .