2013 5th International Conference on Cyber Conflict
K. Podins, J. Stinissen, M. Maybaum (Eds.)

Cyber Deception and Autonomous Attack – Is There a Legal Problem?

William Boothby
Royal Air Force (Ret.)
United Kingdom

Abstract: The publication of the Tallinn Manual on the Law of Cyber Warfare is a huge step forward and now States must decide whether to adopt, formally or otherwise, the rules and guidance it provides. A discussion of deception operations in the cyber age reveals some of the challenges we face in simply transposing existing law of armed conflict rules into cyber terms. Deception operations in warfare are nothing new; some are lawful, and some are not, but does a person have to be deceived for an act that otherwise breaches article 37(1) to be perfidy? How does the law address the improper use of protective indicators and, indeed, espionage in the cyber context? And then we have the crunch question. If cyber deception operations become pervasive so that little or no reliance can be placed, say, on targeting data, what implications does this have for the ability of combatants to comply with distinction, discrimination, proportionality and precautions rules, and does that matter?

Keywords: *Law of Armed Conflict, cyber deception, autonomous attack*

1. INTRODUCTION

The publication this year of the Tallinn Manual[1] has done much to clarify the law on the offensive and defensive use of cyber capabilities in periods of armed conflict. Many matters that were being extensively debated in the literature have been subjected to the critical analysis of the International Group of Experts assembled by the Cooperative Cyber Defence Centre of Excellence in Tallinn. The Experts included *jus in bello* issues in their deliberations and the Manual therefore addresses the rules that regulate the use of cyber force during both international and non-international armed conflicts. While it will be for states to decide whether the conclusions reached by the Experts should guide their warlike activities in the future, there is no doubt that the Manual will at the very least inform the views of States in that regard.

The Experts reached the general conclusion that the law of armed conflict does apply to military cyber operations during and in connection with armed conflicts. Specifically, they reached the clear consensus that the principles of distinction, discrimination, proportionality and the precautionary rules so apply.[2] They also concluded that the rules as to perfidy and ruses of war apply broadly speaking as written in API.[3]

Until a generality of State practice has made the position of States in general clear on particular issues, it will be premature to talk of clear customary law on these cyber warfare issues. Rather what the Manual is doing is to take legal rules that are clearly customary in nature and determine whether there is any apparent reason why they should not apply in cyberspace. The Rules that appear in the Manual are those which, by consensus, the International Group of Experts found to apply in cyberspace as a matter of customary law. The outstanding issue is therefore whether States agree with that interpretation of the Experts.

Until the position of States in that respect becomes clear through practice over coming years and decades, it is sensible to discuss particular issues relating to the conduct of military cyber operations by reference to the black letter Rules and associated Commentaries set forth in the Manual. Those Rules and Commentaries will of course be a valuable resource to States and will assist them to identify perceived gaps in the legal architecture and to determine whether new law is required and, if so, what form it should take. Nevertheless, where there was previously an absence

---

1 Tallinn Manual on the Law of Cyber Warfare, CUP, January 2013.

2 See for example Rules 31, 32, 37 and 49 to 59.

3 Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of International Armed Conflicts, Geneva, 8 June 1977.

of conventional law specifically addressing these subjects, there is now a document asserting contemporary customary and thus universally applicable Rules, and that represents significant progress in this field.

One of the challenges in drafting the Manual was to consider the fundamental differences between cyber activities and more traditional methods of warfare.[4] A recurring issue was to determine whether the existing traditional rules make sense when applied to the cyber domain. Can the peculiarities of cyber operations be accommodated to otherwise well accepted legal rules? To illustrate, and examine the challenges associated with, this issue it is the purpose of the present article to look at one particular aspect of the law on the conduct of cyber hostilities and to consider in some detail the difficulties that the characteristics of cyber operations can be expected to pose.

2. THE ISSUE

Deception is an established and inherently lawful method of undertaking military operations. Perhaps, one of the best known, classical deception operations is Virgil's account in the Aeneid of the use of a wooden model of a horse to infiltrate a Greek unit into the city of Troy after ten years of siege. Another example was the World War Two Operation Mincemeat aimed at convincing the German High Command that the allies would attack Sardinia and Greece in 1943 instead of Sicily, when the means used was the planting of a dead body with false papers concealed upon it to that effect.

So if deception operations are nothing new, we should consider some of the different types of deception operation that have been undertaken in the past. Spaight refers to the use during World War One on occasion of false nationality marks on aircraft. "The inadmissibility of the use of such marks was established, first, by the accusations which the belligerents made against one another of resorting to the practice, secondly by their indignant denials of any complaints that they had done so themselves."[5] On the other hand merely simulating death to avoid being attacked and to permit later escape from a difficult tactical situation has long been seen as legitimate.[6] Spaight also reports the different but relevant and similarly legitimate case of Lieutenant L G Hawker who was seeking to attack a German airship shed at Gontrode in April 1915. It appears that he used "an occupied German

---

4 Throughout the period of the project to produce the Manual, the author was a member of the Group of Experts.

5 J M Spaight, Air Power and War Rights, 3rd Edn (1947) 85-6.

6 Spaight, ibid at page 173.

captive balloon to shield him from fire whilst manoeuvring to drop the bombs".[7] Note also the use of dummy communications to mislead the enemy to believe that fighter aircraft are active when this is not in fact the case[8], also a legitimate practice. Interestingly, Spaight then discusses the legitimacy of tactics during World War One in which an aircraft would simulate landing signals of the enemy's military aircraft in order to enable it to get close to the enemy airfield before dropping its bombs. He concludes that such tactics were lawful because the machine must have been either friend or foe and, in either case, a combat aircraft.[9]

So as can be seen, these deception operations, although they use a considerable variety of techniques, were all aimed at causing the enemy to misunderstand for example the military posture, the identity, the intentions, the manpower capabilities, the resources or the ultimate objectives of the party to the conflict using the deception.

All the indications are that cyber military operations will employ deception to a very considerable degree. Some cyber operations will be so constructed as to appear not to have been undertaken by the State that was in fact responsible. Indeed, in some cases the State undertaking the cyber operation will make it appear that some other State is responsible. In other cases, the cyber operation may be so undertaken as to conceal the very fact of the operation from the enemy.[10] More routinely, damaging cyber packages can be initiated from one computer but may appear to have been sent from an entirely different computer. It may be made to appear that a particular person is the author of a cyber operation when in fact another person originated it. Even if the author of the operation can be identified, false information may be put out to the effect that the originator is acting on behalf of one State or organization when in fact he or she is acting on behalf of another.

Of course these are only examples of the sorts of deception that may be undertaken and it should be borne in mind that multiple deceptions may be used, with the probable intent either that it shall remain permanently unclear who was responsible for a particular event, or that it shall be clear and widely accepted that State or organization A was responsible for it when in fact entity B was in fact answerable. This immediately raises questions over the acceptability of automatic responses to cyber operations. The automatic response may target the computer, system or network where the initial operation appeared to have initiated when, in fact,

---

7 Spaight, ibid, at page 174 citing London Gazette, 8 May 1915.

8 Spaight, ibid, at pages 176-8 cites numerous examples of such ruses in both World Wars.

9 Spaight, ibid at page 179.

10 Consider for example the manner in which the Stuxnet weapon concealed the effect it was having on the Iranian centrifuges from those responsible for monitoring the relevant indicators, thus making it appear that everything was operating normally.

some other computer, system or network was in fact its source. Causing damage to unwilling conduits in this way is likely to prove unacceptable, and risks expanding the scope of conflicts and causing unwanted casualties and damage.

A further question to consider, and the central topic of this paper, is therefore whether this increasing prevalence of military deception that we can foresee as a feature of future military cyber operations is consistent with current interpretations of the law or whether it challenges those interpretations.

3. THE BACKGROUND TO THE CURRENT LAW

We should start our consideration of the law by referring to the Lieber Code. Dr Francis Lieber's text[11] does not have the status of a source of the law[12] but it does indicate what legal thinking was in the mid-nineteenth century on these and related issues. The Lieber Code stipulates that "[m]iliary necessity admits of ….. obstruction of the ways and channels of … communication…. And of such deception as does not involve the breaking of good faith either positively pledged, regarding agreements entered into during the war, or supposed by the modern law of war to exist. Men who take up arms against one another in public war do not cease on this account to be moral beings, responsible to one another and to God."[13] He further noted that military necessity "admits of deception, but disclaims acts of perfidy"[14], a distinction which, as we shall see, lies at the core of the current law.

Dr Lieber included in his text particular provision relating to another form of deception, namely the misuse of a flag of truce. He asserted that if such abuse takes place "for surreptitiously obtaining military knowledge, the bearer of the flag thus abusing his sacred character is deemed a spy" and he goes on to emphasise how necessary the sacred character of the flag of truce is and that its abuse is a heinous crime.[15]

---

[11] Instructions for the Government of Armies of the United States in the Field, 24 April 1863, prepared by Professor Francis Lieber.

[12] This is because the text as such is not one of the law's fundamental principles, nor is it customary law per se and it does not have treaty status; see Statute of the International Court of Justice, article 38.

[13] Lieber Code, article 15.

[14] Lieber Code, article 16.

[15] Lieber Code, article 114. See also Brussels Declaration, article 45, where a parlementaire loses his rights of inviolability if it is shown that he has taken advantage of "his privileged position to provoke or commit an act of treason."

The authors of the Brussels Declaration[16] found a principle of law that has since come to be accepted as one of its cornerstones, namely that the "laws of war do not recognize in belligerents an unlimited power in the adoption of means of injuring the enemy".[17] Applying this principle, they found to be especially forbidden "murder by treachery of individuals belonging to the hostile nation or army"[18] and "making improper use of a flag of truce, of the national flag or of the military insignia and uniform of the enemy, as well as the distinctive badges of the Geneva Convention".[19] The Declaration drew an important distinction between such activities, however, and lawful deception by providing that "ruses of war and the employment of measures necessary for obtaining information about the enemy and the country….are considered permissible".[20]

The Oxford Manual [21] also does not have the status of a source of the law. However, it was written in 1880 by acknowledged experts of the time and it is therefore useful to note that it contained some similar provisions to those in the Brussels Declaration of six years earlier. Article 4 repeats that the means of injuring the enemy are not unlimited, and specifically prohibits perfidious and unjust acts. This Manual requires that conventions, or agreements, between the parties during the conflict must be "scrupulously observed and respected"[22] and that it is forbidden "to make treacherous attempts upon the life of an enemy, as for example by keeping assassins in pay or by feigning to surrender", "to attack an enemy while concealing the distinctive signs of an armed force" or "to make improper use of the national flag, military insignia or uniform of the enemy, of the flag of truce and of the protective signs prescribed by the Geneva Convention"[23].

By the time of the Hague Peace Conferences of 1899 and 1907, thinking and terminology had clarified further. The negotiators included in their texts the general customary admonition that "the right of the belligerents to adopt means of injuring the enemy is not unlimited."[24] More specifically, the Hague Regulations,

---

16 Project of an International Declaration concerning the Laws and Customs of War, Brussels, 27 August 1874.

17 Brussels Declaration, article 12. For the modern formulation see API, article 35(1).

18 Brussels Declaration, article 13(b).

19 Brussels Declaration, article 13(f).

20 Brussels Declaration, article 14, which refers to an exception that the civilian population cannot be forced to take part in military operations against their own country.

21 The Laws of War on Land, Oxford, 9 September 1880

22 Oxford Manual, article 5.

23 Oxford Manual, article 8(b) to (d).

24 For example, Annex to Hague Convention IV Respecting the Laws and Customs of War on Land, art. 22, The Hague, 18 October 1907. Similarly but not identically expressed regulations had been annexed to the Hague Convention II of 1899 but the 1899 text was superseded by the 1907 text and it is therefore on the latter that we will rely.

which have both treaty law and customary status, especially prohibit "kill[ing] or wound[ing] treacherously individuals belonging to the hostile nation or army".[25] Equally importantly, the Regulations made specific provision as to ruses of war, so the distinction that we are discussing in the present article was already embedded in international law in 1899 and 1907. Thus, article 24 of the 1907 text states: "Ruses of war and the employment of measures necessary for obtaining information about the enemy and the country are considered permissible." So an important distinction was made from the outset between acts deceiving the enemy as to matters of protection and lawful ruses and espionage, the latter being accepted as measures in warfare that do not breach the law of war, or in more modern parlance, the law of armed conflict.

4. THE MODERN LAW OF PERFIDY AND RUSES IN API AND THE TALLINN MANUAL

The modern law is to be found in API, article 37. It should be explained at the outset that for the purposes of the present discussion the important distinction is between paragraphs (1) and (2) of that Article, which are as follows:

"(1) It is prohibited to kill, injure or capture an adversary by resort to perfidy. Acts inviting the confidence of an adversary to lead him to believe that he is entitled to, or is obliged to accord, protection under the rules of international law applicable in armed conflict, with intent to betray that confidence, shall constitute perfidy. The following acts are examples of perfidy:

(a) the feigning of an intent to negotiate under a flag of truce or of a surrender;

(b) the feigning of an incapacitation by wounds or sickness;

(c) the feigning of civilian, non-combatant status; and

(d) the feigning of protected status by the use of signs, emblems or uniforms of the United Nations or of neutral or other States not Parties to the conflict.

(2) Ruses of war are not prohibited. Such ruses are acts which are intended to mislead an adversary or to induce him to act recklessly but which infringe no rule of international law applicable in armed conflict and which are not perfidious because they do not invite the confidence of an adversary with respect to protection under the law. The following are examples of such ruses: the use of camouflage, decoys, mock operations and misinformation."

---

[25] Note that the word 'treachery' means for all practical purposes the same as the word 'perfidy' as used in the following discussion of the modern law; Tallinn Manual, paragraph 1 of the commentary accompanying rule 60.

Both of these paragraphs are, it must be appreciated, concerned with deception operations. However, some such operations are lawful under paragraph (2) whereas others are rendered unlawful by paragraph (1). It is therefore important to identify the critical points of difference between the two classes of operation. Both classes of operation invite the confidence of the enemy in relation to matters which are in fact untrue. Both classes of operation are aimed at persuading the enemy either to act or to refrain from acting on the basis of that induced false appreciation of the facts.

The critical point of difference is the nature of the false belief that is being induced in the enemy by the operation. In the second, lawful class of deception operations, the deception does not breach a rule of the law of armed conflict and is not inviting confidence "with respect to protection under the law". In the first, unlawful class of deception operation, the deception is directed at inducing the adverse party to believe that there is either a right to, or a duty to give, protection under the law. AMW notes that a "typical example of perfidy would be to open fire upon an unsuspecting enemy after having displayed the flag of truce, thereby inducing the enemy to lower his guard".[26]

The Tallinn Manual expresses the perfidy rule in terms that are very similar to article 37. The significant point of difference is that the Manual's rule[27] omits reference to capture as a result of the perfidy, simply because the customary rule does not extend to capture, unlike the rule in article 37. [28] It should be emphasized that an act of perfidy that does not result in death, injury or capture does not breach either rule. The Tallinn Manual makes the useful point that the person deceived need not necessarily be the same person as the one whose death or injury results provided that the person who is killed or injured is in fact the intended target of the attack.[29] However, the perfidy must be the proximate cause of the death or injury. While there may be a time delay between the two events, it is causal proximity that is relevant here.[30]

---

[26] AMW, commentary accompanying Rule 111(a), paragraph 8.

[27] Tallinn Manual, rule 60.

[28] See paragraph 2 of the Commentary associated with Rule 60, which notes that Hague Regulations article 23(b) makes no mention of capture. It is noted there that the corresponding war crime under the Rome Statute of the International Criminal Court, 1998, namely article 8(2)(b)(xi) in relation to international armed conflicts, also makes no mention of capture resulting from perfidy. The corresponding war crime in the Rome Statute that arises in relation to non-international armed conflicts is in article 8(2)(e)(ix) which refers to "[k]illing or wounding treacherously a combatant adversary". Notwithstanding the omission of capture from the Rome Statute war crimes, the ICRC contends that the customary law of armed conflict rule includes capture; ICRC Customary Humanitarian Law Study, Rule 65 and commentary at page 225 where it is suggested that the consequences of capture may not be grave enough to constitute the act of perfidy a war crime.

[29] Tallinn Manual, commentary accompanying Rule 60, paragraph 4.

[30] Tallinn Manual, commentary accompanying Rule 60, paragraph 6.

The Group of Experts then considered a situation which brings us rather closer to the topic of the present paper. They considered whether a person has to be deceived for the perfidy rule to be broken or whether the rule extends to deception of a machine. The example referred to in the Manual is that of a cyber deception operation that targets a pacemaker fitted, for example, to an enemy commander, causing the pacemaker to malfunction thus killing the commander. If the cyber operation betrays the confidence of the computer controlling the pacemaker, a majority of the Experts concluded that the perfidy rule is broken. The minority view was that for perfidy to be made out, the deception must operate on a human mind in the prohibited way.[31]

5. THE MODERN LAW OF PERFIDY AND RUSES IN OTHER MANUALS

The Air and Missile Manual[32] finds a rule expressed in similar terms to article 37(1) of API.[33] Importantly, perfidious action that results in damage but not in death, injury or capture does not constitute a breach of the law of armed conflict and, by extension, does not amount to a war crime.[34]

AMW then notes, most importantly in relation to the current discussion, that the mere fact "that a person is fighting in civilian clothing does not constitute perfidy"[35] although the same Manual notes that the person fighting in this way may not be entitled to combatant immunity and may thus be prosecuted and punished under domestic law.[36]

AMW also finds a rule as to ruses of war the effect of which largely reflects the rule as expressed in article 37(2) of API, but which employs somewhat different language.[37] It notes that the fact that the ruse results in death, injury or capture of personnel of the adverse party does not per se cause the attack to be prohibited as

---

[31] See Tallinn Manual, commentary accompanying Rule 60, paragraph 9.

[32] Program on Humanitarian Policy and Conflict Research, Harvard University, Manual on International Law Applicable to Air and Missile Warfare, published with a commentary in March 2010 and referred to collectively here as 'AMW'.

[33] AMW, Rule 111(a) and (b). Note also the US Commanders' Handbook on the Law of Naval Operations, NWP 1-14, paragraph 12.12 states a rule in similar language.

[34] AMW, commentary accompanying Rule 111(a), paragraph 7.

[35] AMW, commentary accompanying Rule 111(b), paragraph 4.

[36] The same paragraph of the AMW Commentary cites a useful example of perfidy, where the individual advances to an advantageous position "under the cover of being a civilian in order to fire on, and kill or injure, an unsuspecting enemy".

[37] The differences in language do not seem to produce significant difference in intended meaning.

perfidy provided that deception as to protected status is not involved.[38] It suggests as examples of lawful ruses the following activities in air warfare, namely "(a) mock operations[39]; (b) disinformation[40]; (c) false military codes and false electronic, optical or acoustic means to deceive the enemy (provided that they do not consist of distress signals, do not include protected codes, and do not convey the wrong impression of surrender)[41]; (d) use of decoys and dummy-construction of aircraft and hangars; and (e) use of camouflage".[42]

The same Manual also gives examples of air operations that would constitute perfidious conduct. The listed examples are "(a) the feigning of the status of a protected medical aircraft, in particular by the use of the distinctive emblem or other means of identification reserved for medical aircraft; (b) the feigning of the status of a civilian aircraft; (c) the feigning of the status of a neutral aircraft; (d) the feigning of another protected status; and (e) the feigning of surrender."[43]

Highly significantly from the perspective of the present text, whether or not such behavior is perfidious, AMW finds the following conduct is always prohibited, namely improper use by aircraft of distress codes, signals or frequencies and use of any aircraft which is not a military aircraft as a means of attack.[44] Improper use in this regard means any use outside normal purposes. So, for example, distress signals must be reserved for their humanitarian purposes,[45] and any military use of such signals that is outside the scope of humanitarian activity and which is, say, aimed at facilitating the undertaking of an attack, would be prohibited by the Rule. There is a fine distinction to be considered here. Thus, if a pilot of an aircraft sends a false distress signal that will clearly breach the Rule. If, however, the same pilot refrains from sending such a signal, but so flies his aircraft as to cause those on the

---

38 AMW, commentary accompanying Rule 113, paragraph 3.

39 AMW cites as examples air attacks on the Pas de Calais during the weeks leading up to D-Day in 1944 or the movement of, e.g., an aircraft carrier to create a false impression as to the likely nature of an attack. Similarly, simulated air attacks may be undertaken, as lawful ruses of war, as a device to persuade the enemy to activate its air defences and thus provide valuable targeting information. The common theme here is the presentation to the enemy of a false picture of what is occurring.

40 AMW gives as an example an attempt to induce the enemy to surrender by creating the false impression that he is surrounded, or that an overwhelming attack is about to occur; AMW, commentary accompanying Rule 116(b), paragraph 2, where the distinction is noted between such lawful activities and the use of false information as to civilian, neutral or other protected status which would not be lawful; ibid, paragraph 3.

41 The use of enemy IFF codes, or the use of the enemy's password to avoid being attacked when summoned by an enemy sentry or inducing a false return on the enemy radar screen indicating the approach of a larger force than is the case are all cited in AMW as lawful ruses; commentary accompanying Rule 116(c), paragraphs 2 and 3.

42 AMW, Rule 116.

43 AMW, Rule 114.

44 AMW, Rule 115. Distress codes signals and frequencies do not for these purposes include IFF codes; AMW commentary accompanying Rule 115(a), paragraph 5.

45 AMW, commentary accompanying Rule 115(a), paragraph 1.

ground to form the incorrect view that the aircraft has been damaged, that would not breach the rule.[46] Contrast the circumstance discussed in paragraph 4 of the same commentary, namely where the pilot of an aircraft simulates a situation of distress with the purpose of creating the false impression that personnel deploying from the aircraft by parachute are entitled to protection under article 42 of API.[47] In these circumstances, if the deploying personnel are in fact paratroopers "this could amount to prohibited perfidy if it leads to the killing, injuring (or capturing) of an adversary."[48]

The UK Manual gives the following examples of ruses: "transmitting bogus signal messages and sending bogus despatches and newspapers with a view to their being intercepted by the enemy; making use of the enemy's signals, passwords, radio code signs, and words of command; conducting a false military exercise on the radio while substantial troop movements are taking place on the ground; pretending to communicate with troops or reinforcements which do not exist…; and giving false ground signals to enable airborne personnel or supplies to be dropped in a hostile area, or to induce aircraft to land in a hostile area".[49]

6. IMPROPER USE OF CERTAIN INDICATORS

The other deception-related provisions of API that we will discuss in the present paper are to be found in articles 38 and 39 and relate to the misuse of the emblems specified in those Articles. Thus, article 38(1) prohibits making "improper use" of the red cross or red crescent[50] or of other emblems, signs or signals provided for in the Conventions or in the Protocol and further prohibits the deliberate misuse of other internationally recognized protective emblems, signs or signals.[51] Importantly, Article 9 of Annex I to API, as amended on 30 November 1993, addresses means of electronic identification of medical transports.

---

[46] Note in this regard that a damaged aircraft is not necessarily a disabled aircraft, neither is it necessarily a surrendering aircraft; consider the discussion at AMW, commentary accompanying Rule 115(a), paragraph 3.

[47] "No person parachuting from an aircraft in distress shall be made the object of attack during his descent."; API, art. 42(1).

[48] AMW, commentary accompanying Rule 115(a), paragraph 4 and note API, art. 42(3): "Airborne troops are not protected by this Article."

[49] U.K. Manual, para. 5.17.2.

[50] Additional Protocol III to the Geneva Convention, article 2(1), applies the same prohibition to the Red Crystal adopted by that Instrument also as a distinctive emblem.

[51] As the Tallinn Manual notes at paragraph 2 of the Commentary accompanying Rule 62, this would extend to the distinctive sign for cultural property, for civil defence, the flag of truce and the electronic protective markings set out in Annex I to API; Cultural Property Convention, articles 16 and 17, API, art. 66, Hague Regulations, art. 23(f) and API, Annex I, paragraph 9. See also AMW, Rule 112(a) and (b).

As the Tallinn Manual makes clear, these are absolute provisions that do not require death, injury or capture as an essential ingredient of a breach while the term 'improper use' is considered to comprise any use other than that for which the emblem, sign or signal was intended.[52] Accordingly, this and the following examples of improper use of emblems etc are prohibited irrespective of whether the acts concerned also amount to perfidy.[53] It will be noted from paragraphs 6 and 7 of the commentary accompanying Rule 62 in the Tallinn Manual that the Group of Experts was divided as to whether the rule is specifically restricted to misuse of the emblem, sign or signal as such or whether misuse of a domain name such as 'icrc.org' to like effect would also be prohibited. The author takes the provisional view that the former interpretation is *lex lata* while the latter view may reflect *lex ferenda*.

While the focus in those provisions is on 'improper use', in article 38(2) "[i]t is prohibited to make use of the distinctive emblem of the United Nations, except as authorised by that Organization".[54] While it is clear that the prohibition will extend to unauthorized use of the emblem by electronic means, the same division of opinion as described in the previous paragraph applies to whether breach of the rule requires use of the emblem as such.[55]

Any use of flags, insignia or military emblems of the enemy is prohibited "while engaging in attacks or in order to shield, favour, protect or impede military operations".[56] The Tallinn Manual adds the words "while visible to the enemy" to the rule, to reflect the majority view among the experts that "it is only when the attacker's use is apparent to the enemy that the act benefits the attacker or places its opponent at a disadvantage".[57] However, where the use of the enemy's emblem in cyber communications is concerned, the Tallinn Manual is explicit, opining "it is permissible to feign enemy authorship of a cyber communication", basing this view on State practice regarding lawful ruses.[58]

---

[52] Tallinn Manual, commentary accompanying Rule 62, paragraphs 3 and 4, citing in the latter instance the ICRC Study, commentary accompanying Rule 61.

[53] AMW, chapeau to Rule 112 and commentary accompanying Rule 111(a), paragraph 9.

[54] For the application of this rule in cyber operations, see Tallinn Manual, Rule 63, citing NWP 1-14, paragraph 12.4, the UK Manual paragraph 5.10.c and the AMW Manual, Rule 112(e).

[55] See commentary accompanying Rule 63, Tallinn Manual. It will be appreciated that if the United Nations becomes a party to an armed conflict, its military personnel who are combatants and the objects it uses to make an effective contribution to the hostilities will be lawful targets. Misuse of its emblem by an adverse Party to such a conflict would, in those circumstances, amount to improper use of an enemy emblem as opposed to misuse of the United Nations emblem; AMW, commentary accompanying Rule 112(e), paragraph 3.

[56] API, article 39(2), AMW Rule 112(c) and Tallinn Manual, Rule 64.

[57] Tallinn Manual, commentary accompanying Rule 64, paragraph 4.

[58] Citing the extract from the UK Manual noted earlier in the present paper.

Article 39(1) of API prohibits making use of "the flags or military emblems, insignia or uniforms of neutral or other States not Parties to the conflict" and the Tallinn Manual finds, subject to the traditional rules of naval warfare,[59] a customary rule expressed in identical terms.[60] Any such use is unlawful, so the word 'improper' is not included in Article 39(1) nor in the corresponding Rule in the Manuals. There was however, as the Tallinn Manual explains, division among the Experts as to whether the use of other indicators, such as the domain name of the neutral's Ministry of Defence, would constitute a breach of the rule in circumstances in which the emblem as such is not employed.[61]

7. ESPIONAGE

As AMW notes, "espionage consists of activities by spies", adding, perhaps rather more usefully, that "a spy is any person who, acting clandestinely or on false pretences, obtains or endeavours to obtain information of military value in territory controlled by the enemy, with the intention of communicating it to the opposing Party."[62] Rule 66(a)of the Tallinn Manual makes it clear that cyber espionage and other forms of intelligence gathering directed at an adverse party to the conflict do not breach the law of armed conflict.[63]

The Tallinn Manual describes as 'clandestine' acts that are undertaken secretly or secretively, whereas the term 'under false pretenses' refers to acts so conducted as to create the impression that the individual has the right to access the information concerned.[64] Importantly, a person who obtains information about an adversary while the information gatherer is located outside enemy controlled territory is

---

[59] See API, art. 39(3).

[60] Tallinn Manual, Rule 65; see also AMW, Rule 112(d).

[61] Tallinn Manual, commentary accompanying Rule 65, paragraph 4.

[62] AMW, Rule 118. Article 29 of the Hague Regulations of 1899 and 1907 provided: "An individual can only be considered a spy if, acting clandestinely or, on false pretenses, he obtains, or seeks to obtain information in the zone of operations of a belligerent, with the intention of communicating it to the hostile party. Thus, soldiers not wearing a disguise who have penetrated into the zone of operations of the hostile army, for the purpose of obtaining information, are not considered spies…." Note the Lieber Code stipulated that a "spy is a person who secretly, in disguise or under false pretense, seeks information with the intention of communicating it to the enemy"; Lieber Code, article 88 and see Brussels Declaration, article 19.

[63] Tallinn Manual, Rule 66(a), AMW, Rule 119. However, a combatant who acts as a spy loses the right to be a POW and may be treated as a spy if captured before he reaches the army on which he depends; Tallinn Manual, Rule 66(b).

[64] Tallinn Manual, commentary accompanying Rule 66, paragraph 2 citing API Commentary, paragraph 1779. See also AMW, Rule 120 where it is noted that an individual is not engaged in espionage if, while gathering the information, he is in the uniform of his armed forces. However, members of military aircrew who wear civilian clothes inside a properly marked military aircraft are not spies; AMW, commentary accompanying Rule 120, paragraph 2.

not engaged in espionage. In the cyber context, therefore, most acts of remotely undertaken information gathering will not constitute espionage, whereas close access cyber operations to obtain information from a targeted closed computer system using, for example, a memory stick will be espionage if the targeted computer is located within the enemy's zone of operations provided that the other elements of espionage are present.[65]

8. DO FORESEEABLE NOTIONS OF CYBER OPERATIONS CHALLENGE THE LAW?

We have now seen how the law regulates deception operations as they have been undertaken during traditional types of military operation. The question that now needs to be considered is whether the pervasive use of cyber deception to which we referred in the first section of this paper has implications for these traditional legal rules. To put it more succinctly, will these new kinds of cyber operation, and the associated extensive deception operations, challenge the law by requiring that existing legal rules be adjusted to permit such deceptions to be used more frequently, or will the existing rules prevail, for example because only deceptions that comply with traditional interpretations of the law will in fact be permitted and, thus, undertaken?

To make sense of this generalized question, we should very briefly consider a number of scenarios. They are purely illustrative, do not reflect all foreseeable kinds of cyber deception operation that may be relevant, but will at least give an indication as to the sorts of legal issue that may be expected to arise. The scenarios are:

A State A undertakes a remote access cyber attack making it appear that the attack has been undertaken by State B. State B is a co-belligerent of State A, so there is no breach of international law by State A as a result of the deception *per se*. However, if the subject of the cyber attack were to mount an automatic response attack against State B, this would likely breach international law as being an unlawful use of force or, even, an armed attack. This implies a need for caution to be exercised before responding, to seek to ensure that the true author of the initial attack is being engaged.

---

[65] If undertaken by a civilian, remote cyber information gathering and close access cyber espionage are likely to constitute direct participation in the hostilities, which, if undertaken by a civilian, would render him or her liable to attack while so engaged. It is also likely to breach the domestic law of the territory where the activity occurs and the persons concerned are therefore liable to be tried for the relevant offences; AMW, Rule 121.

B During an international armed conflict, State A undertakes a remote access cyber attack using a worm incorporated within an attachment to an Email and making it appear that the attack has been undertaken by State B. State B is a neutral and a reproduction of its national flag is employed to make the Email appear to have come from an authentic State B source. Making use of the neutral's flag clearly breaches API, article 39(1). If the flag were not to be used and the deception were based on use of the neutral government's domain name, such as '.gov.uk', the Tallinn Manual's Experts were divided as to whether such activity is unlawful.

Again, however, an automatic response against State B would, on the face of it, constitute an unlawful use of force or armed attack, and in both this and the previous example, it would seem advisable that diplomatic activity be undertaken to seek to confirm responsibility for the attack before a use of force, cyber or otherwise, in response is decided upon.

C A false Email sent to enemy personnel causes them to believe that they are being invited to attend a meeting to discuss the surrender of the unit sending the Email. The sending unit has no intention of surrendering, but the deceived personnel suffer a road accident on the way to the proposed meeting resulting in death or injury. The deception operation is not, arguably, the proximate cause of the accident and while the message was perfidious, the Rule is not broken because the death or injury are not proximately caused by the perfidy.

D A false Email sent to enemy personnel causes them to believe that they are being invited to attend a command group meeting. The Email appears, falsely, to have been sent by the Enemy superior commander to his subordinate commanders. As it is permissible to feign enemy authorship of cyber communications, the operation would appear to be lawful, even if enemy personnel are as a result killed or injured.

E Having hacked into the enemy computerized Common Operating Picture programme, false data is inserted making it appear that friendly forces are concentrated distant from their true location. This would be a lawful ruse.

If the false data were to make it appear that friendly forces are concentrated in or near a small town populated with civilians, and if the enemy as a result attacks that location causing incidental damage and casualties among the civilians, the perfidy rule would, arguably, not have been breached because no civilian status has been feigned in relation to the friendly forces themselves.

F Personnel from a State that is not party to API who are members of a military unit pretend to have civilian status. They dress in civilian clothes, alter the unit's website so as to make it appear civilian, include assertions in the website of its civilian status and omit all references to military ranks in any electronic communications from the unit.

On the approach of an attacking unit, the personnel from the State not party to the conflict are not attacked because of their apparently civilian status, but after that attack, succeed in capturing enemy personnel and in damaging their military equipment. No perfidious offence is committed as the customary perfidy provision does not extend to capture, and neither the customary nor the API rule extends to damage caused by the perfidy.

G A cyber operation deceives the targeted computerized perimeter security system to believe that enemy personnel are in fact friendly forces. The enemy personnel then enter the closed military facility protected by the security system and wreck the facility, capture its personnel and kill the commander. If the attackers enter in uniform, the operation would not be prohibited perfidy. If they enter in civilian clothes, it likely would be.

H A cyber operation deceives the targeted computer that protects the perimeter of a military, closed IT system. The deception causes the protecting computer to believe that an attachment to an Email has been received from a non-threatening civilian source and, thus, may be opened in accordance with IT protocols without undertaking certain preliminary checks. The attachment, when opened, causes the server to which the targeted computer is connected to shut down thus denying service to all users, with the result that the targeted unit's water purification system instantly malfunctions causing death and injury through disease/infections. According to the majority view among the Group of Experts, this deception of the targeted computer would be perfidy and, as it leads to death and injury, would be prohibited.

9. CONCLUSION

It is clear that deception operations will become of increasing importance as cyber warfare techniques become more widely employed in armed conflict. The traditional rules draw a vital distinction between lawful deception, and that which is prohibited because, causing death, injury or capture, it leads the adversary to believe that he is entitled to or is obliged to accord legal protected status. There is no reason to believe that this traditional distinction will be either eroded or abandoned in the cyber context. The focus in the definition of espionage on the geographical location of the spy may seem outdated in an age when remote access cyber operations may be employed to intrude into the most secret, protected and sensitive parts of the enemy's information architecture. Outdated or not, the geographical element in the espionage definition is customary, and thus binds all States, and seems unlikely to change.

The rules prohibiting the use of certain flags, emblems, insignia or uniforms also may appear to some to be somewhat anachronistic. The degree to which the capabilities of, and risks posed by, cyber operations will call the adequacy of these rules into question remains to be seen. For the time being at least, they have stood the test of time and are consistent, essentially, with the philosophy underpinning the perfidy rule.

Having put forward this case in support of the legal status quo, the author would offer one word of caution. It is this. If increasingly pervasive cyber capabilities are so used that deception operations become the rule rather than, relatively speaking, the exception, and if as a result little or no reliance can in future be placed on the information that would traditionally support targeting decision making, what are the consequences for the practical ability of combatants to comply with the distinction, discrimination, proportionality and precautions rules that lie at the core of targeting law? It seems to the author that some at least concrete basis for reliable decision making is central to the practical delivery of these protective principles. Widespread use of deception must not, it is suggested, become the cause of a slide into ‘anything goes’.